

# Une méthode alternative pour caractériser l'environnement « obésogénique » de l'enfant

## Pertinence d'une analyse factorielle des correspondances multiples (AFCM)

**Benjamin C. Guinhouya, Geoffroy K. Apété, Djamel Zitouni, Mohamed Lemdani, Christian Vilhelm, Alain Durocher, Hervé Hubert**

DANS **SANTÉ PUBLIQUE** 2010/2 Vol. 22 , PAGES 165 À 179  
ÉDITIONS **S.F.S.P.**

ISSN 0995-3914

DOI 10.3917/spub.102.0165

Date de mise en ligne : 01/05/2010

Article disponible en ligne à l'adresse

<https://stm.cairn.info/revue-sante-publique-2010-2-page-165?lang=fr>



Découvrir le sommaire de ce numéro, suivre la revue par email, s'abonner...  
Scannez ce QR Code pour accéder à la page de ce numéro sur Cairn.info.



**Distribution électronique Cairn.info pour S.F.S.P..**

Vous avez l'autorisation de reproduire cet article dans les limites des conditions d'utilisation de Cairn.info ou, le cas échéant, des conditions générales de la licence souscrite par votre établissement. Détails et conditions sur [cairn.info/copyright](http://cairn.info/copyright).

Sauf dispositions légales contraires, les usages numériques à des fins pédagogiques des présentes ressources sont soumises à l'autorisation de l'Éditeur ou, le cas échéant, de l'organisme de gestion collective habilité à cet effet. Il en est ainsi notamment en France avec le CFC qui est l'organisme agréé en la matière.

# Une méthode alternative pour caractériser l'environnement « obésogénique » de l'enfant. Pertinence d'une analyse factorielle des correspondances multiples (AFCM)

*An alternative method to describe and "obesogenic" environment for children*

**Benjamin C. Guinhouya** <sup>(1)</sup>, **Géoffroy K. Apété** <sup>(1)</sup>, **Djamel Zitouni** <sup>(2)</sup>, **Mohamed Lemdani** <sup>(2)</sup>, **Christian Vilhelm** <sup>(1)</sup>, **Alain Durocher** <sup>(1)</sup>, **Hervé Hubert** <sup>(1)</sup>

**Résumé :** Le but de cette étude est de caractériser l'environnement « obésogénique » d'un groupe d'enfants d'âge scolaire en utilisant une analyse factorielle des correspondances multiples (AFCM), comme approche alternative aux choix méthodologiques traditionnels, applicable, même aux échantillons de taille modeste. Quatre-vingt-onze enfants (39 filles et 52 garçons) âgés de  $10,0 \pm 0,9$  ans ont été aléatoirement recrutés au sein de 2 établissements scolaires publics français. Les données concernant leur cadre familial, l'implication des parents, leur temps télévisuel et leur comportement alimentaire ont été obtenues par questionnaire. Leur niveau d'activité physique et leur temps sédentaire ont été évalués au moyen d'un accéléromètre (MTI Actigraph, modèle 7164) pendant 3 jours, y compris un jour férié. Les données ont été traitées au moyen d'une AFCM. Une méthode d'estimation du risque a ensuite été appliquée en considérant la distribution des enfants suivant leur statut vis-à-vis du surpoids/obésité dans le plan factoriel produit par l'AFCM. Les caractéristiques « obésogéniques » sont apparues comme 4 possibilités de combinaison entre l'environnement familial et des comportements variés vis-à-vis de l'activité physique, la sédentarité et l'alimentation. Le risque relatif (RR) de surpoids/obésité était de 2,64 [1,52 ; 4,57] ( $p < 0,0001$ ) pour une combinaison entre un environnement familial « défavorisé » + une association activité physique faible et alimentation riche en matières grasses et de 0,36 [0,14 ; 0,94] ( $p < 0,05$ ) pour un environnement familial « favorisé » + une association activité physique forte et alimentation pauvre en matières grasses. L'AFCM apparaît suffisamment robuste et pertinente pour orienter efficacement les hypothèses étiologiques et les décisions concernant les stratégies d'intervention individuelle ou collective.

**Mots-clés :** Obésité - analyse factorielle - environnement - comportements.

(1) Institut Lillois d'Ingénierie de la Santé (ILIS), Lab. Santé Publique, EA 2694. UDSL, Université Lille-Nord-de-France, 42, rue Ambroise-Paré, 59120 Loos.

(2) Faculté des Sciences Biologiques et Pharmaceutiques, Département de Biomathématiques, EA2694. UDSL, Université de Lille-Nord-de-France, 3, rue du Prof.-Laguesse, 59006 Lille cedex.

**Summary:** The purpose of this study was to describe an “obesogenic” environment for a group of schoolchildren using a multiple correspondence analysis (MCA) as an alternative approach to traditional methodological choices. MCA is applicable even for small samples. Ninety-one children (39 girls and 52 boys) aged  $10.0 \pm 0.9$  years were randomly recruited from two French public schools. Data on their family context, parental involvement, television time and their eating habits were obtained through questionnaires. Their level of physical activity and sedentary time were assessed using an accelerometer (MTI Actigraph model 7164) for three days, including a holiday. The data were processed using an MCA together with a technique for estimating relative risks (RRs) of overweight/obesity according to the distribution of children in the factorial plane produced by the MCA. The “obesogenic” factors appeared as four possible combinations between family environments and various behaviours with regard to physical activity, sedentary behaviour and diet. The RR of overweight/obesity was 2.64 [1.52, 4.57] ( $P < 0.0001$ ) for a combined association of a “disadvantaged” family environment + low physical activity and high fat diet. The RR of overweight/obesity was 0.36 [0.14, 0.94] ( $P < 0.05$ ) for an association of a “privileged” family environment + high physical activity and low fat diet. Thus, MCA appears sufficiently robust and relevant to effectively guide etiologic hypotheses and decisions about individual and collective intervention strategies.

**Keywords:** obesity - factor analysis - environment - behaviours.

## Introduction

Malgré la prévalence élevée du surpoids/obésité dans la population pédiatrique, ses causes spécifiques demeurent mal expliquées [3]. Elles sont nombreuses, multifformes et probablement interactives. Pour expliquer les augmentations du surpoids/obésité durant la dernière décennie dans la population pédiatrique, certains auteurs soutiennent l'idée d'une contribution prédominante de facteurs modifiables liés à l'environnement et aux comportements [11]. En revanche, d'autres données semblent suggérer une contribution biologique et génétique prépondérante, en raison notamment de la multiplication des accouplements assortis (« *assortive mating* ») [4]. Si des travaux intensifs restent nécessaires pour identifier les facteurs causaux principaux de l'obésité infantile, la description et le contrôle des facteurs comportementaux et environnementaux potentiellement impliqués dans ce problème de santé paraissent essentiels dans la perspective de l'infléchir. Ceci est particulièrement relevant que l'influence des gènes de susceptibilité pourrait être améliorée ou diminuée grâce à l'adoption des comportements appropriés [4].

Ainsi, le cadre de vie de l'enfant interagirait d'une manière très spécifique avec son comportement pour déterminer son statut pondéral. Cependant, il apparaît primordial de mieux décrire le rôle, l'importance relative et les interactions entre différentes caractéristiques de l'environnement (*e.g.*, le cadre familial) et du comportement de l'enfant et son excès pondéral. En regard de la complexité évidente de ces relations [7], les méthodes d'analyse courantes ne permettent pas de les appréhender dans leur globalité, sans hypothèses aprioristes, fondamentalement simplificatrices et réductrices de la réalité. En l'occurrence, ou les auteurs se limitaient bien souvent à des analyses univariées dont la portée reste limitée, ou dans le meilleur des cas, il s'agissait d'utiliser une analyse multivariée hypothético-déductive laissant place à de nombreuses hypothèses. Dans cette perspective, la régression logistique a correspondu à la méthode d'analyse traditionnelle [2]. La régression logistique est l'une des méthodes multivariées explicatives les

plus utilisées en épidémiologie et en santé publique, où elle permet de prédire les facteurs qui caractérisent un groupe de sujets malades par rapport à des sujets sains. Le plus grand des attraits du modèle logistique est qu'il permet d'estimer des rapports de cotes (RC) ou « *odds ratios* (OR) » ; une approximation du risque relatif (RR) lorsque la prévalence de la pathologie est faible. Cette approximation perd de son sens lorsqu'on étudie par exemple le surpoids/obésité des jeunes qui a tendance à affecter plus de 10 %, et jusqu'au-delà 30 % de certaines populations. L'utilisation d'une régression logistique implique également de respecter un certain nombre de contraintes qu'il n'est pas toujours aisé d'atteindre dans certaines études : indépendance entre les variables explicatives ; degré de liberté important (*i.e.* une taille d'échantillon relativement élevée) ; normalité de la distribution des variables quantitatives intégrées dans le modèle ou encore la nécessité d'une relation forte entre chacune des variables quantitatives et la variable dépendante [2]. Enfin, le formalisme de la loi logistique répond à la démarche commune des tests d'hypothèses probabilistes, dont le corollaire est l'acceptation ou le rejet d'une hypothèse dans le cas d'une adéquation ou non entre les données et le modèle logistique. Cette démarche qui implique de fixer un seuil de significativité statistique, permet de limiter la marge de subjectivité qui peut être associée à l'analyse de la complexité. Cette approche probabiliste autorise également une certaine générabilité des résultats obtenus. Néanmoins, cette démarche essentiellement stochastique peut trouver quelques limitations au regard de l'explication biologique ou clinique de certains phénomènes [1].

Bien qu'elle soit classée comme une méthode d'analyse descriptive ou exploratoire, l'analyse factorielle des correspondances multiples (AFCM) – qui n'impose aucune restriction ni sur la distribution des variables, ni sur la nécessité d'une quelconque linéarité – semble être une approche alternative adaptée à l'étude de la complexité en santé et pour construire des « *modèles qui répondent aux données et non l'inverse* » [18]. Cette méthode part donc d'un principe opposé à celui des tests d'hypothèses fondés sur l'adéquation des données à un formalisme pré-admis. L'utilisation de l'AFCM n'impose donc aucune distribution probabiliste des données. En conséquence, les résultats qu'on en obtient doivent être transposés avec à-propos à la population dont est issu l'échantillon d'étude [10]. Mais, l'AFCM tire tout son intérêt de la prise en compte simultanée (et sans présomption sur le caractère confondant ou interactif des variables considérées) d'un nombre relativement important de variables pour ne pas négliger à tort des relations qui pourraient exister. De ce fait, l'AFCM peut s'accommoder de variables interdépendantes alors qu'une collinéarité importante entre les variables quantitatives explicatives peut poser des problèmes d'estimation des paramètres par la loi logistique. Cette possibilité, qui permet par l'AFCM d'appréhender avec une certaine finesse la communauté des caractéristiques qui réunit un groupe d'individus, peut toutefois poser un problème de redondance des variables sélectionnées lors de l'interprétation des résultats. La principale limitation liée à l'utilisation de l'AFCM concerne le risque de « subjectivité » dans l'interprétation des dimensions, que seules la pratique, l'expérience et une profonde connaissance du problème de santé étudié permettent d'atténuer. En effet, la simplicité apparente des représentations géométriques et la compréhension des données numériques associées ne

doivent, à aucun moment, éloigner de la complexité inhérente à la fois à cette méthode et au problème étudié. Néanmoins, ce type d'analyse permet une description exhaustive des tendances et des typologies, offrant ainsi une base de travail propice pour l'orientation des actions de santé publique, la prédiction et la décision politique.

Le but de cette étude est donc de tester l'intérêt de l'AFCM, comme une option méthodologique applicable, même aux échantillons de taille modeste, pour l'exploration des caractéristiques « obésogéniques » de l'enfant.

## Population et méthodes

### Participants et schéma d'étude

Cent soixante-six enfants des 2 sexes, âgés de 7 à 12 ans, inscrits en cycle III dans deux établissements scolaires primaires publics – en milieu urbain – de l'académie de Lille (année scolaire 2005-2006) ont été invités à participer à cette étude d'observation (Figure 1).

De la cohorte éligible, 136 élèves (57 filles et 79 garçons) ont été recrutés sur une base de volontariat, et 80 % des enfants de la cohorte initiale (*i.e.*, 110 enfants) ont été aléatoirement choisis pour la présente étude, en raison du nombre d'accéléromètres disponibles pour évaluer leurs comportements de mouvement. Les enfants sélectionnés n'ont rapporté aucun problème de santé lié à la pratique d'activités physiques et participaient aux séances d'éducation physique et sportive à l'école. Cette étude transversale à visée analytique a reçu l'accord du Comité consultatif de Protection des Personnes (CPP) de Lille. Avant le début des études, les participants (enfants, parents et tuteurs légaux) ont reçu des informations verbales et écrites sur le détail du protocole et les risques éventuels encourus par les enfants. Pour être admis, parents/tuteurs légaux et enfants ont remis un consentement écrit.

## Mesures

### Paramètres anthropométriques

La taille a été évaluée au moyen d'une toise murale à 0,1 cm près. La stature complète est définie comme étant la distance maximale entre le sol et le vertex de la tête. Les enfants devaient se tenir droit, les pieds joints, la tête droite avant la prise de la mesure. La masse corporelle a été obtenue à l'aide d'une balance électronique. La mesure de la masse corporelle était réalisée à 0,1 kg près. Deux lectures ont été faites pour chaque paramètre anthropométrique. Une troisième mesure était requise si les deux premières différaient de 0,4 cm pour la taille et de 0,4 kg pour la masse corporelle. La valeur retenue correspondait à la moyenne des deux premières mesures, sinon à la valeur médiane des 3 mesures. Pendant les mesures, les enfants étaient habillés de façon légère, les pieds nus.

L'indice de masse corporelle (IMC) est calculé comme le rapport de la masse corporelle au carré de la taille (exprimée en mètres). Il a servi à classer les enfants suivant leur statut pondéral (normopondéral vs. surpoids/

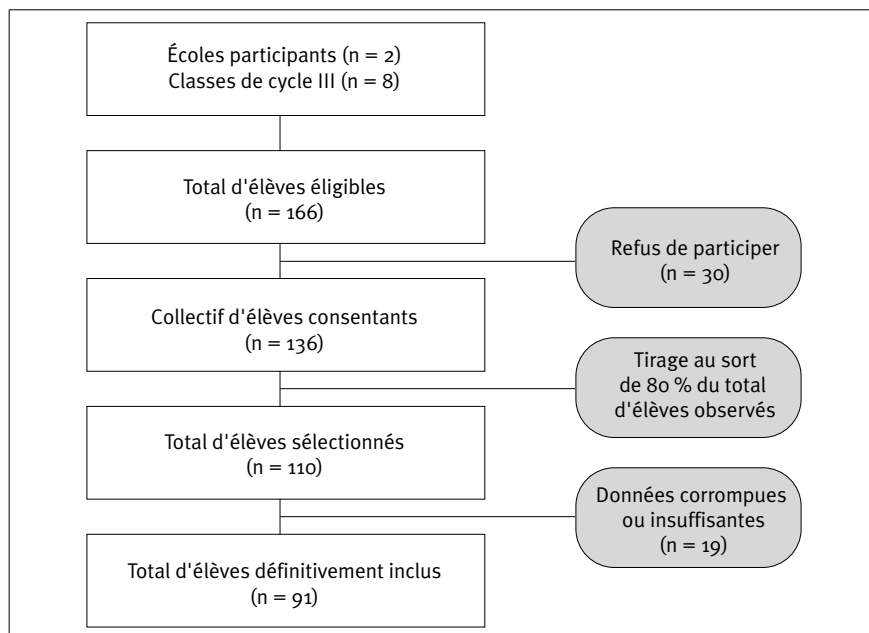


Figure 1 : Diagramme du recrutement des élèves.

obésité) en utilisant les références de l'*International Obesity Task Force* (IOTF, [6]).

### Environnement familial

Les éléments du contexte familial des enfants ont été recueillis à l'aide d'un questionnaire rempli par les parents. Ces données concernaient les possibilités économiques du ménage, mais également la faculté des parents à inciter et à participer à une vie familiale active. La classification sociale et économique des familles a été réalisée en regroupant l'activité professionnelle des parents en 2 classes : classe défavorisée (*e.g.*, professions manuelles : ouvrier(ère)s, employé(e)s, chômeurs et femmes au foyer si la profession du conjoint n'a pas été renseignée) et classe favorisée (*e.g.*, professions non manuelles et intellectuelles). Il importe de préciser que la profession d'un seul des parents avait été considérée puisque, la plupart d'entre eux (75 %) n'avaient fourni aucune indication sur la profession de leur conjoint. Des informations concernant le type d'habitation, l'habitat de l'enfant ou encore le moyen de déplacement principal de la famille ont été acquises par le biais de ce questionnaire. Le type d'habitat a permis de définir 2 modalités selon que la famille est propriétaire de son logement (*Propriétaire*) ou en location (*Locataire*). Deux modalités ont été retenues concernant le logement des enfants, selon qu'ils partagent ou non la même chambre que leur fratrie.

La structure familiale définie comme le nombre de personnes composant le ménage a également été obtenue par le questionnaire. Le nombre de

personnes présentes dans la famille a été dichotomisé en fonction de la valeur médiane obtenue sur cet échantillon (5 personnes). Aussi pouvait-on distinguer les familles de 5 individus ou plus ( $\geq 5$  *mbrs*) des familles de moins de 5 personnes ( $< 5$  *mbrs*).

Par ailleurs, deux modalités afférentes à la participation ou non des enfants à un loisir institutionnel et à la fréquentation d'un centre de loisirs le mercredi ont été obtenues. Enfin, les habitudes familiales vis-à-vis de l'activité physique avaient été rapportées par les parents. Il s'agissait pour les parents d'indiquer la fréquence (régulièrement, occasionnellement, rarement) des occasions d'activité physique (promenades à pied, à vélo, jogging) réalisées en famille avec les enfants.

### Comportement alimentaire

Les habitudes alimentaires ont été évaluées au moyen d'un questionnaire de fréquence alimentaire rempli en classe avec l'aide des enseignants. Ce questionnaire ne permettait pas d'estimer la valeur calorique des aliments mais d'obtenir une estimation « brute » des matières grasses contenues dans l'alimentation de l'enfant. Le questionnaire comprend une vingtaine d'items alimentaires généralement consommés par les enfants. Chaque enfant devait répondre à la question « *combien de fois manges-tu chacun des aliments suivants ?* », en cochant une proposition parmi : « *pas beaucoup* », « *un peu* » et « *beaucoup* ». La moyenne des scores est utilisée comme l'indicateur du contenu en matières grasses de l'alimentation. Plus cette moyenne est élevée plus l'alimentation habituelle de l'enfant est considérée comme riche en matières grasses. Les propriétés psychométriques de la version originale de ce questionnaire portant sur 83 items alimentaires montrent que ce questionnaire des habitudes alimentaires est fiable et valide par rapport au « *multiple 24-h recalls* » [23]. La présente version est une adaptation française de celle utilisée chez des enfants américains [16]. Une analyse de fiabilité sur des données obtenues après 48 heures auprès d'un sous-groupe de 39 enfants, choisis aléatoirement parmi le total d'enfants observés, a révélé un coefficient de corrélation intraclasse pour cohérence absolue de 0,89 à 95 % IC : [0,78-0,94]. Les valeurs obtenues ont été dichotomisées suivant le score médian de 1,81 de cet échantillon. Ainsi, on pouvait distinguer les enfants ayant des préférences alimentaires orientées vers les matières grasses ( $\geq 1,81$  : *Gras*) de ceux dont l'alimentation est moins riche en matières grasses ( $< 1,81$  : *Peu gras*).

### Temps passé devant un écran

Le temps passé par les enfants devant un écran (télévision, vidéo, ordinateur, consoles) a été considéré comme une construction sédentaire spécifique. Il a été obtenu au moyen du questionnaire rempli par les parents. Les valeurs obtenues pour les jours ouvrables et fériés ont été dichotomisées pour distinguer, les enfants qui passent moins de 14 h/semaine devant un écran de ceux qui passent plus de 14 h/semaine. Cette catégorisation est basée sur les travaux de Castebon et Rolland-Cachera [5] qui ont montré que seulement 2 % à 3 % d'enfants français passaient plus de 4 h/jour devant l'écran.

### Activité physique et sédentarité

L'activité physique et la sédentarité ont été évaluées objectivement à l'aide d'un accéléromètre du type MTI ActiGraph (Model AM 7164, FL, USA) calibré pour collecter les informations toutes les minutes pendant une semaine. Il a été recommandé aux enfants et notifié aux parents, que l'appareil ne devait pas être porté lors d'activités de baignade (*e.g.*, douche, bain ou piscine). L'actigraphe ou moniteur d'activité physique est une sonde destinée à enregistrer l'activité physique des individus pour des variations d'accélération d'amplitude allant de 0,05 à 2 G. L'appareil mesure  $5 \times 4 \times 1,5$  cm et pèse 35 gr. Son alimentation est assurée par une pile en lithium. L'appareil a été conçu avec un capteur dont le contenu fréquentiel permet de détecter toute la bande de fréquences comprises entre 0,25 et 2,5 Hz. Les valeurs codées, enregistrées à la vitesse de 10 signaux/seconde, sont converties en une unité arbitraire, le « *count* » dont l'importance reflète l'amplitude des signaux reçus à la fin d'une période cyclique prédéfinie par l'expérimentateur. À l'issue de la période cyclique, l'intégrateur numérique est raffûté, alors que l'information recueillie est préalablement stockée en mémoire.

#### Réduction des données actigraphiques

Les enfants ont en moyenne porté l'actigraphe pendant 3,6 jours. Les valeurs obtenues entre 7 h 00 et 21 h 00 pendant 3 jours y compris un jour férié (mercredi ou dimanche, avec un ratio de 85 % mercredi/15 % dimanche) ont servi aux analyses. Les valeurs brutes ont été rangées suivant le jour, la date et l'heure des enregistrements. Pour chaque jour, une inspection visuelle des valeurs a permis de savoir si un minimum de 10 heures effectives d'enregistrement étaient opérées, en considérant comme valeurs aberrantes ou un manque de respect de la procédure, toute suite de 20 valeurs consécutives égales à « 0 ». Tous les enfants inclus dans l'analyse finale ont produit des données  $\geq 10$  h d'enregistrement. Dix-neuf enfants de la cohorte sélectionnée ont été exclus en raison de données corrompues ou insuffisantes sur le nombre de jours ou d'heures d'enregistrement.

Différents paramètres associés à ces mesures ont été calculés : le volume d'activité physique ( $AP_{tot}$ ) correspond à la valeur moyenne (exprimée en counts par min [cpm]) obtenue dans une journée. Le temps passé à réaliser des activités d'intensité modérée à vigoureuse (APMV exprimé en min/jr) est calculé en utilisant le seuil de Puyau *et al.* [19]. Le temps sédentaire (Sed, exprimé en min/jr) est obtenu en utilisant le seuil de Reilly *et al.* [20]. Ce seuil permet de classer les enfants suivant leur temps sédentaire et non sédentaire avec une sensibilité de 83 % et une spécificité de 82 % par rapport à l'observation comportementale, utilisée comme méthode de référence [20]. Ces données sur les temps actifs et sédentaires ont servi à créer des catégories (*e.g.*, actifs vs. non-actifs ; sédentaires vs. non-sédentaires).

Le volume d'activité physique quotidien des enfants ( $AP_{tot}$ ) a été utilisé pour répartir les enfants en 3 groupes : un groupe faiblement actif ( $AP_{tot} < 454$  cpm), un groupe moyennement actif ( $455 \leq AP_{tot} \leq 540$  cpm) et un groupe fortement actif ( $AP_{tot} > 541$  cpm).

En raison du faible temps passé dans une APMV, cette variable n'a pu qu'être dichotomisée en utilisant la valeur de 20 min·j<sup>-1</sup> comme référence.

Ainsi, les enfants ayant passé plus de 20 minutes par jour au-dessus de 3 200 cpm sont considérés comme actifs (« *moitié active* ») ; sinon, ils sont considérés comme peu actifs (« *moitié inactive* »). Le temps sédentaire des enfants (Sed) a permis de regrouper les enfants en 3 groupes : les enfants faiblement sédentaires (Sed < 681 min·j<sup>-1</sup>), les enfants moyennement sédentaires (682 min·j<sup>-1</sup> ≤ Sed ≤ 712 min·j<sup>-1</sup>) et les enfants fortement sédentaires (Sed > 713 min·j<sup>-1</sup>).

## Analyses statistiques

### Analyses univariées

Les valeurs sont présentées sous la forme de moyenne ± écart-type ou de fréquence suivant la nature des variables. Un test de Kolomogorov-Smirnov complété avec la correction de Lilliefors a permis de vérifier la normalité des données. Dans le cadre des analyses univariées, des comparaisons entre garçons et filles ont été réalisées à l'aide d'un test « *t* » de Student pour échantillons indépendants. Le cas échéant, un test « *U* » de Mann et Whitney a été utilisé. Des relations entre variables quantitatives ont été recherchées au moyen d'un coefficient de corrélation de Pearson/de Spearman, selon le cas. D'autres dépendances ont été examinées à l'aide d'un test du  $\chi^2$ . Dans tous les cas, le seuil de significativité était fixé à  $P < 0.05$ .

### Analyses multivariées

L'ensemble des variables caractérisant le comportement des enfants, le cadre et les habitudes de vie de la famille a été traité par une analyse factorielle des correspondances multiples (AFCM). Cette méthode d'analyse de données de type exploratoire se destine à la réduction du nombre d'informations contenues dans un tableau dit de « grande taille », en établissant les correspondances nécessaires. L'AFCM est appropriée pour l'examen de la structure des données indépendamment d'un modèle aprioriste et des probabilités de distribution. Ainsi, à partir de données interactives trop nombreuses pour être appréhendées directement, cette technique permet d'extraire les tendances les plus marquantes, les hiérarchise et élimine les effets marginaux ou ponctuels qui perturbent la perception globale des faits. L'AFCM autorise des représentations géométriques interprétables des individus et des caractères et s'apparente, sur le principe à l'analyse en composante principale (ACP) sans rotation, mais s'en distingue par le fait qu'elle est basée sur des écarts pondérés. Cette particularité donne la même importance relative aux écarts à l'indépendance quel que soit le nombre de cas. Une telle approche globale des phénomènes est beaucoup plus riche d'enseignements que l'examen séparé des différents facteurs en présence. Le résultat se présente sous la forme d'un nombre restreint de dimensions suffisant pour exprimer l'essentiel du phénomène étudié. L'analyse graphique de l'existence de groupes de variables se réalise dans l'espace défini par les dimensions retenues. Elle s'opère en considérant la proximité entre les variables, mais également la séparation des modalités, parce que plus les variables sont proches plus elles sont interreliées, et

celles séparées par des quadrants correspondent à des groupes de profils opposés. Dans la présente étude, les 2 premières dimensions ont été conservées pour l'analyse. En nous appuyant sur la répartition des enfants, suivant leur statut pondéral, dans le plan défini par ces 2 dimensions (représentant les caractéristiques « obésogéniques »), nous avons pu estimer le risque relatif (RR) du surpoids/obésité des enfants de cet échantillon. Le RR a été défini comme le rapport entre l'incidence du surpoids/obésité chez les sujets exposés à un agrégat de caractéristiques données et l'incidence du surpoids/obésité chez les sujets restants (non exposés aux caractéristiques précitées). L'intervalle de confiance à 95 % (IC 95 %) de ce risque a été construit au moyen de la méthode dite « semi-exacte » :

$$IC_{95\%}(RR) = RR \times \exp \left[ \pm 1,96 \sqrt{\left( \frac{b}{a \times E_1} \right) + \left( \frac{d}{c \times E_0} \right)} \right], \text{ avec}$$

$a$  = nombre d'enfants en surpoids/obésité exposés aux caractéristiques « obésogènes » ;  
 $b$  = nombre d'enfants normo-pondéraux exposés aux caractéristiques « obésogènes » ;  
 $c$  = nombre d'enfants en surpoids/obésité non-exposés aux caractéristiques « obésogènes » ;  
 $d$  = nombre d'enfants normo-pondéraux non-exposés aux caractéristiques « obésogènes » ;  $E_0$  = nombre total d'enfants non-exposés aux caractéristiques « obésogènes » ;  $E_1$  = nombre total d'enfants exposés aux caractéristiques « obésogènes ».

L'ensemble des analyses a été réalisé avec le logiciel SPSS 13.0 (SPSS Inc., Chicago USA).

## Résultats

### Caractéristiques générales de la population

Les données concernant ~67 % de la population initiale, soit 91 élèves (39 filles et 52 garçons) âgés en moyenne de  $10,0 \pm 0,9$  ans, ont été utilisées dans les analyses. En termes de classification économique et sociale, on peut noter que 34 % des enfants appartiennent à la classe « défavorisée » et 66 % à la classe « favorisée ». Comme l'indique le Tableau I, il n'y avait pas de différence statistiquement significative entre les garçons et les filles de cet échantillon par rapport aux paramètres anthropométriques. En revanche, les garçons sont apparus plus actifs que les filles pour le volume d'activité physique et le temps passé dans une APMV ( $P < 0,0001$ ). De même, les garçons ont un temps sédentaire significativement moins important que les filles ( $P < 0,05$ ).

### Environnement « obésogénique » et risque de surpoids/obésité

Le total d'inertie expliquée par le nuage de points défini par les 2 dimensions retenues est de 43,4 % (dimension 1 : 26,2 % et dimension 2 : 17,2 %). La répartition des variables dans le plan délimité par ces 2 dimensions permet d'entrevoir une organisation en 4 grands ensembles opposés 2 à 2 (Figure 2a). Dans les quadrants A et A1 (Figure 2a), on observe une agrégation de variables relatives au cadre de vie des enfants. Le quadrant A caractérise un regroupement de modalités associées à l'appartenance à une catégorie professionnelle et sociale défavorisée, à un

**Tableau I : Descriptif général de la population**

	Fille (n = 39)	Garçon (n = 52)	Groupe (n = 91)
Age (an)	10,1 ± 0,8	9,9 ± 0,9	10,0 ± 0,9
Taille (m)	1,42 ± 0,08	1,41 ± 0,07	1,41 ± 0,07
Masse (kg)	38,3 ± 8,4	36,4 ± 7,9	37,2 ± 8,2
IMC (kg·m <sup>-2</sup> )	18,7 ± 3,0	18,1 ± 2,5	18,4 ± 2,8
Surpoids/Obésité (références IOTF)	36 %	31 %	33 %
AP <sub>tot</sub> (cpm)	498,5 ± 118,0	630,9 ± 193,2*	574,4 ± 177,2
APMV (min·j <sup>-1</sup> )	20,7 ± 11,9	38,8 ± 24,7**	31,1 ± 22,0
Sed (min·j <sup>-1</sup> )	720 ± 34	701 ± 39*	709 ± 38

\*\* Différence statistiquement significative entre les filles et les garçons,  $P < 0,0001$

\* Différence statistiquement significative entre les filles et les garçons,  $P < 0,05$

IOTF : International Obesity Task Force

AP<sub>tot</sub> : Volume d'activité physique réalisé par jour

APMV : Activité physique d'intensité modérée à vigoureuse

Sed : Temps passé dans des activités sédentaires

nombre élevé de personnes vivant dans un habitat loué et avec des enfants partageant une chambre commune. Autour de cet agrégat, on observe une proximité de la non participation des enfants à des loisirs institutionnels et de la rareté des occasions d'activité physique en famille. Par ailleurs, les enfants semblent passer plus de 14 hr/semaine devant un écran. Dans le quadrant A1 positionné en opposition du quadrant A, ressortent des caractéristiques opposées à celles qui viennent d'être énumérées.

En ce qui concerne l'espace délimité par les quadrants B et B1 (Figure 2a), on pourra noter une proximité des variables reflétant les comportements des enfants vis-à-vis de l'activité physique et de l'alimentation. Le quadrant B fournit des indications sur un fort niveau d'activité physique (aussi bien pour le volume d'activité physique « AP fort » que le temps passé à réaliser une APMV « Plus actif ») associé à un niveau de sédentarité faible (« Sed faible »). Les habitudes alimentaires privilégient les aliments pauvres en matières grasses. À l'inverse, dans le quadrant B1, on peut remarquer un agrégat de comportements, qu'on peut qualifier, de défavorables vis-à-vis de l'activité physique et de l'alimentation.

La Figure 2b présente une répartition des enfants dans le plan délimité par les 2 dimensions retenues. Son analyse doit s'appuyer sur la Figure 2a en considérant les interactions entre les quadrants A → B1, A → B, A1 → B, et A1 → B1 à l'origine de la formation des espaces (E1), (E2), (E3) et (E4) respectivement.

Ces espaces seront considérés comme l'environnement « obésogénique » et traduisent les interactions entre un contexte familial spécifique et des comportements individuels variés. Compte tenu de la distribution des enfants selon leur statut pondéral dans ces espaces, nous obtenons les risques relatifs (RR) suivants, récapitulés dans la Figure 3.

- E1 : environnement familial « défavorisé » + association activité physique faible et alimentation riche en matières grasses : RR = 2,64 ; 95 % IC = [1,52 ; 4,57] ( $P = 0,001$ ).

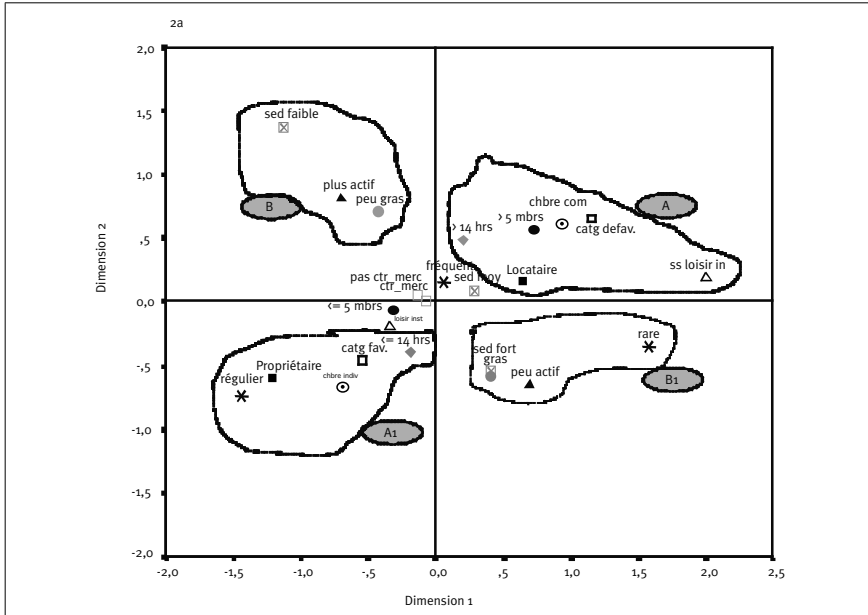


Figure 2a : Quantification et coordonnées factorielles des différentes modalités

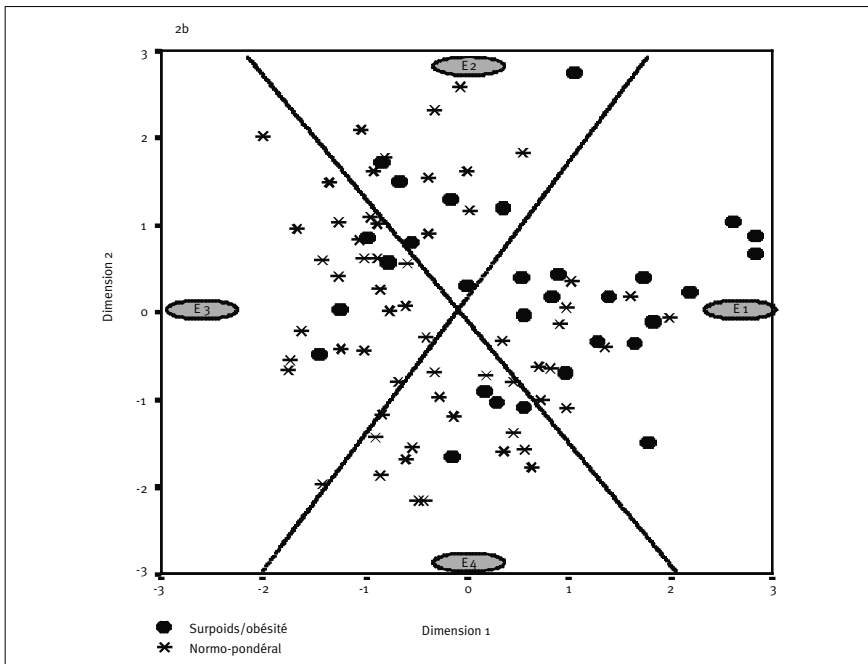


Figure 2b : Répartition des enfants dans le plan factoriel en fonction de leur statut vis-à-vis du surpoids/obésité

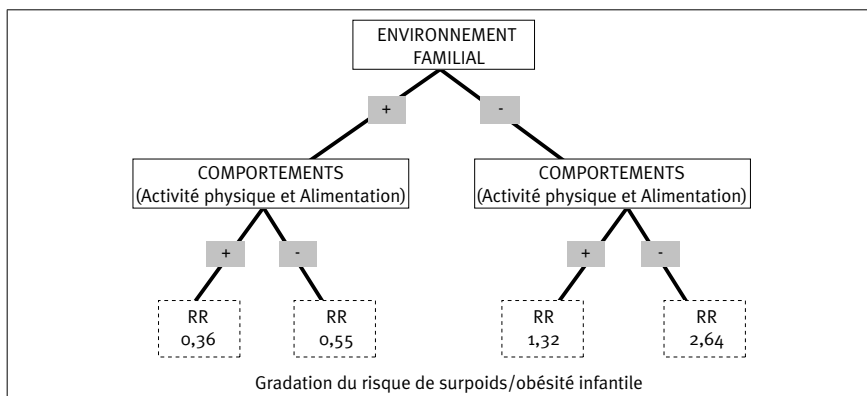


Figure 3 : Arbre logique structurant les interactions entre les caractéristiques de l'environnement familial et les comportements individuels et les niveaux de risque de surpoids/obésité qui en résultent

- E2 : environnement familial « défavorisé » + association activité physique forte et alimentation pauvre en matières grasses : RR = 1,32 ; 95 % IC = [0,68 ; 2,57] ( $P = 0,56$ )
- E3 : environnement familial « favorisé » + association activité physique forte et alimentation pauvre en matières grasses : RR = 0,36 ; 95 % IC = [0,14 ; 0,94] ( $P = 0,03$ )
- E4 : environnement familial « favorisé » + association activité physique faible et alimentation riche en matières grasses : RR = 0,55 ; 95 % IC = [0,22 ; 1,38] ( $P = 0,19$ ).

## Discussion

Cette étude avait pour objectif de tester une nouvelle approche méthodologique dans l'exploration des interrelations entre les variables environnementales et comportementales potentiellement impliquées dans le surpoids/obésité de l'enfant, sur un échantillon de taille modeste. Il s'agissait d'identifier la typologie des interactions entre les facteurs de l'environnement « obésogénique » en utilisant une méthode d'analyse factorielle. La complémentation de la démarche avec une technique d'estimation du risque de surpoids/obésité infantile en a sans doute accru l'intérêt et l'originalité. Les résultats indiquent des liens entre le surpoids/obésité infantile et l'environnement « obésogénique », avec une gradation du risque en fonction des caractéristiques « obésogéniques » envisagées.

À notre connaissance, cette étude est la première à utiliser une analyse factorielle telle que l'AFCM pour structurer les interrelations entre des « cofacteurs » environnementaux et comportementaux du surpoids/obésité de l'enfant. Nos observations sont donc difficilement comparables aux résultats issus d'analyses classiques. Certains profils obtenus dans la présente étude sont concordants avec la littérature existante, notamment en regard de l'influence du statut économique et social [12, 15, 21, 22] ou du rôle de l'alimentation [13, 14] et de l'activité physique [8, 17]. Toutefois, la

diversité des profils tels que rapportés dans le présent travail est une de ses caractéristiques particulières. La différenciation entre les 2 groupes de variables (facteurs liés au contexte familial vs. comportement des enfants) qui forment le plan factoriel reflète le fait que les comportements des enfants ne sont pas complètement figés par le contexte social et économique de leur cadre familial. Cette observation permet d'apporter quelques nuances aux croyances habituelles de l'impact négatif d'un milieu socio-économique « défavorisé » sur les comportements santé, et de considérer au contraire les variations de profils résultant d'interactions nombreuses. La méthode apparaît d'autant plus robuste que les résultats ont été obtenus avec un degré de liberté faible par rapport aux études épidémiologiques classiques. Une observation qui demeure toutefois intrigante dans notre analyse concerne l'agonisme systématique des comportements alimentaires avec l'activité physique. En effet, on observe que lorsque le comportement de l'enfant vis-à-vis de l'activité physique est « positif », les choix alimentaires s'orientent préférentiellement vers une faible consommation de matières grasses (comportement alimentaire « positif ») et inversement. Une hypothèse consisterait à dire qu'une certaine forme de réceptivité s'est mise en place grâce aux campagnes médiatiques, notamment télévisuelles, sur l'alimentation et l'activité physique. On pourrait aussi arguer que les enfants ont rapporté des habitudes alimentaires sous l'influence de facteurs tels que la désirabilité sociale. En effet, compte tenu de la nature auto-déclarative du questionnaire alimentaire, on ne peut exclure la possibilité d'une sous-estimation des consommations alimentaires [9]. Certaines données soutiennent l'existence de ce type de biais, en particulier chez les jeunes en surpoids/obésité [24]. Mais, ce résultat peut aussi être simplement le fait d'un biais d'information directement lié au questionnaire alimentaire. Cet outil initialement développé à destination des enfants américains [23], requiert sans doute une meilleure adaptation et une validation par rapport aux habitudes culturelles françaises.

D'un point de vue pratique, si nos résultats peuvent paraître redondants par rapport à la littérature actuelle, cette étude met en exergue la pertinence des méthodes d'analyse factorielle peu utilisées dans les études épidémiologiques et en santé publique. Elle démontre notamment la capacité de ces méthodes à traiter la complexité des problèmes de santé, notamment dans le cas d'événements multifactoriels comme le surpoids/obésité infantile. L'AFCM offre sans doute un cadre de travail qui mêle des données quantifiées à une interprétation à la fois souple et didactique, en proposant une structuration appropriée de l'information, base d'analyses étiologiques intéressantes. Dans la présente étude, l'AFCM a simplifié la mise en évidence des caractéristiques « obésogéniques » des enfants, et oriente vers quelques pistes d'individualisation des actions de prévention ou d'affinement des stratégies de promotion des comportements santé.

Il faudrait tout de même tenir compte de la nature transversale de notre étude, fondée essentiellement sur une cohorte d'enfants dont la surcharge pondérale est déjà constituée. Nos observations ne permettent donc pas d'établir des relations de causalité avec les déterminants précoces de l'obésité infantile. Par ailleurs, nos résultats soulèvent quelques questions concernant les coordonnées centrales. En effet, comme le montre la

figure 2a, un nombre important de modalités a des coordonnées proches de (0,0) ; ce qui réduit leur pouvoir discriminant. Cette observation implique d'une part que, les sujets ayant participé à cette étude sont très ressemblants pour un grand nombre de caractéristiques et d'autre part, l'existence probable d'un biais de sélection. L'aspect volontaire de la participation des écoles et des enfants en est peut-être un élément explicatif. Même si les données utilisées dans cette analyse proviennent d'une base de données un peu plus large, la manière dont le recrutement initial avait été abordé ne peut garantir l'absence de ce type de biais. La validité interne de nos résultats mériterait donc d'être appréciée à partir d'un échantillon de taille plus importante. Enfin, il aurait été également enrichissant pour notre approche de disposer d'informations supplémentaires sur les enfants (*e.g.*, auto-efficacité, compétence nutritionnelle perçue), les parents (*e.g.*, situation maritale, statut vis-à-vis de l'obésité), et l'environnement physique dans lequel vivent les familles. L'ensemble de ces points mériterait d'être pris en compte dans les études ultérieures, en vue d'en tirer des enseignements plus généraux. Ces études doivent également proposer et intégrer une forme de « *cross-validation* » du (des) modèle(s) retenu(s).

Ce travail dont l'objet principal était de tester la pertinence d'une analyse factorielle des correspondances multiples (AFCM) en épidémiologie, offre un nouveau cadre méthodologique et quelques perspectives pour l'analyse des relations souvent complexes entre les facteurs environnementaux et la santé. L'exemple du surpoids/obésité infantile suggère l'applicabilité de l'AFCM – même sur des échantillons de taille modeste – et sa pertinence à orienter efficacement les hypothèses étiologiques et les décisions concernant les stratégies d'intervention tant au niveau individuel que collectif.

## REMERCIEMENTS

Nous tenons à remercier les enfants et leurs parents ainsi que les enseignants des écoles Taine et Chopin impliqués dans cette étude pour leur bonne volonté. Nos remerciements s'adressent également au service de promotion de la santé de la ville de Villeneuve d'Ascq, qui, par l'entremise de Yveline Redlich et Catherine Allogio, a encouragé et contribué à la réalisation de ce travail.

## BIBLIOGRAPHIE

1. Akobeng AK. Confidence intervals and p-values in clinical decision making. *Acta Paediatr* 2008;97:1004-7.
2. Aminot I, Damon MN. Régression logistique : intérêt dans l'analyse de données relatives aux pratiques médicales. *Rev Med Ass Maladie* 2002;33:137-43.
3. Bar-Or O. Juvenile obesity. Is school-based enhanced physical activity relevant?. *Arch Pediatr Adolesc Med* 2005;159:996-7.
4. Bouchard C. Childhood obesity: are genetic differences involved. *Am J Clin Nutr* 2009;89 (suppl):15-8S.
5. Castetbon K, Rolland-Cachera MF. Surpoids et obésité chez les enfants de 7 à 9 ans. Paris: CNAM (<http://lesrapports.ladocumentationfrancaise.fr/BRP/054000194/0000.pdf>); 2000.
6. Cole TJ, Bellizzi MC, Flegal KM, Dietz WH. Establishing a standard definition of childhood obesity worldwide: international survey. *Br Med J* 2000;320:1-6.
7. Corvol P. L'homme et son environnement ou de la complexité dans le domaine de la santé. *Med/Sci* 2000;16:1015-6.
8. Dencker M, Thorsson O, Karlsson MK, Linden C, Eiberg S, Wollmer P, et al. Daily physical activity related to body fat in children aged 8-11 years. *J Pediatr* 2006;149:38-42.

9. Deschamps V, De Lauzon-Guillain B, Lafay L, Borys JM, Charles MA, Romon M. Reproducibility and relative validity of a food-frequency questionnaire among French adults and adolescents. *Eur J Clin Nutr* 2009;63:282-91.
  10. Greenacre MJ. Theory and applications of correspondence analysis. London: Academic Press 1984.
  11. Hill JO, Wyatt HR, Reed GW, Peters JC. Obesity and the environment: where do we go from here? *Science* 2003;299:583-855.
  12. Lioret S, Maire B, Volatier JL, Charles MA. Child overweight in France and its relationship with physical activity, sedentary behaviour and socioeconomic status. *Eur J Clin Nutr* 2007;61:509-16.
  13. Lioret S, Touvier M, Dubuisson C, Dufour A, Calamassi-Tran G, Lafay L, et al. Trends in Child Overweight Rates and Energy Intake in France From 1999 to 2007: Relationships With Socioeconomic Status. *Obesity (Silver Spring)* 2009;17:1092-100.
  14. Lioret S, Touvier M, Lafay L, Volatier JL, Maire B. Are eating occasions and their energy content related to child overweight and socioeconomic status? *Obesity (Silver Spring)* 2008;16:2518-23.
  15. Lioret S, Touvier M, Lafay L, Volatier JL, Maire B. Dietary and physical activity patterns in French children are related to overweight and socioeconomic status. *J Nutr* 2008;138:101-7.
  16. McMurray RG, Harrell JS, Bangdiwala SI, Deng S. Cardiovascular disease risk factors and obesity of rural and urban elementary school children. *J Rur Health* 1999;15:365-74.
  17. Ness AR, Leary SD, Mattocks C, Blair SN, Reilly JJ, Wells J, et al. Objectively measured physical activity and fat mass in a large cohort of children. *PLoS Med* 2007;4:e97.
  18. Panagiotakos DB, Pitsavos C. Interpretation of epidemiological data using multiple correspondence analysis and log-linear models. *J Data Science* 2004;2:75-86.
  19. Puyau MR, Adolph AL, Vohra FA, Butte NF. Validation and calibration of physical activity monitors in children. *Obesity Res* 2002;10:150-7.
  20. Reilly JJ, Coyle J, Kelly L, Burke G, Grant S, Paton JY. An objective method of measurement of sedentary behavior in 3- to 4-year olds. *Obes Res* 2003;11:1155-8.
  21. Rolland-Cachera MF, Bellisle F. No correlation between adiposity and food intake: why are working class children fatter? *Am J Clin Nutr* 1986;44:779-87.
  22. Romon M, Duhamel A, Collinet N, Weill J. Influence of social class on time trends in BMI distribution in 5-year-old French children from 1989 to 1999. *Int J Obes* 2005;29:54-9.
  23. Speck BJ, Bradley CB, Harrell J, Belyea MJ. A food questionnaire for youth: psychometric analysis and summary of eating in adolescents. *J Adolesc Health* 2001;28:16-25.
  24. Ventura AK, Loken E, Mitchell DC, Smiciklas-Wright H, Birch LL. Understanding reporting bias in the dietary recall data of 11-year-old girls. *Obesity* 2006;14:1073-84.
-