

LES FACTEURS DE RISQUE EN EPIDEMIOLOGIE ANIMALE

F. MADECO^[1], Christine FOURICHON^[2]

RESUME : La recherche des facteurs de risque doit ouvrir la voie à des possibilités de prévention des troubles de santé. Différentes méthodes d'obtention des facteurs de risque ont été proposées. Lorsqu'il s'agit de l'étude de phénomènes complexes, une voie complémentaire à celle des tests classiques consiste à utiliser les méthodes descriptives de l'analyse des données. Cette voie présente cependant nombre de contraintes relatives à la collecte des données et à la conduite des opérations de traitement statistique.

SUMMARY : Research related to risk-factors is an essential sector in epidemiology since it opens towards opportunities for prevention. Several methods have been proposed with the scope to obtain the risk-factors. When complex diseases are concerned (syndroms), an alternative way to the conventional tests (RR, odds-ratio...) is to use multivariate descriptive methods such as the correspondence analysis and the cluster analysis. This option has nevertheless certain rules and constraints that have to be respected as well during data collection as during data processing.

*
* *

I. INTRODUCTION

Dans le langage courant, le mot "risque" est généralement associé à l'existence d'un danger, d'un inconvénient possible. La connaissance de l'existence de risques conduit alors les individus à se prémunir. Les exemples sont légion dans notre vie quotidienne : "aux risques et périls de l'acheteur", "assurances tous risques"... Chez l'épidémiologiste et le statisticien, on retrouve derrière le mot risque cette idée qu'un événement peut survenir. Celui-ci peut être souhaité et favorable (on parle également de chance) ou défavorable. Ainsi, existe-t-il, sous-jacente, l'idée d'association de conditions permettant l'apparition ou non du phénomène. Cette apparition peut être aléatoire et donc imprévisible. Inversement, elle peut correspondre à la mise en oeuvre de phénomènes mesurables, l'ensemble répondant à certaines règles. L'un des objectifs des travaux de l'épidémiologiste est précisément de découvrir ces phénomènes, de les mesurer puis de trouver les règles ou, plus généralement, les circonstances qui président à leur apparition. On pressent ainsi que les calculs vont concerner les relations statistiques et les probabilités. C'est ainsi que les "facteurs de risque" d'un phénomène donné seront appréhendés. On en déduira des "populations à risque", des "zones à risque", etc. La présente publication se propose d'évoquer les grandes caractéristiques des maladies en élevage intensif, puis de dégager les lignes directrices des démarches possibles permettant d'obtenir les facteurs de risque.

* Texte de la conférence présentée le 31 mai 1990.

[1] CNEVA, Station de Pathologie Porcine BP 53 - 22440 Ploufragan.

[2] ENV Nantes, Route de Gachet CP 3013 - 44087 Nantes cedex 03.

II. RELATION DE CAUSALITE ET FACTEURS DE RISQUE EN PATHOLOGIE ANIMALE

On associe habituellement une maladie à un agent spécifique, qu'il soit microbien ou non. Le vocable utilisé pour désigner les problèmes de santé est parfois révélateur (salmonellose, pasteurellose, fusariotoxicose, intoxication au mercure...). Inversement, des agents infectieux reçoivent une appellation se rapportant aux manifestations associées à leur intervention : pestivirus, Herpès virus... L'association fréquente entre un agent et des manifestations pathologiques présente de nombreux avantages mais elle tend également à se transformer rapidement, dans l'esprit des gens, en une relation de causalité. Or, la relation de causalité suppose des conditions bien particulières que nombre d'appellations ne vérifient pas (Evans, 1978) (tableau I).

Tableau I : Conditions permettant à un facteur associé à une maladie d'en être considéré comme une cause (Evans, 1978).

-
1. L'incidence et/ou la prévalence de la maladie doivent être supérieures chez les individus exposés au facteur.
 2. L'exposition au facteur doit être plus fréquente chez les sujets malades.
 3. L'exposition doit précéder l'apparition de la maladie.
 4. Il doit y avoir un spectre de réponses perceptibles et mesurables chez l'hôte mis en présence du facteur (apparition d'anticorps...).
 5. L'élimination du facteur doit s'accompagner d'une réduction de l'incidence de la maladie.
 6. Toute intervention visant à empêcher ou à modifier la réponse de l'hôte doit réduire le degré d'expression de la maladie.
 7. La maladie doit pouvoir être reproduite expérimentalement.
-

Tout comme la relation de causalité, le concept de "risque" trouve également une large place en épidémiologie. Devant certaines dérives dans l'utilisation et l'interprétation du mot, une approche rationnelle a été souhaitée (Grundy, 1973). Cet auteur souligne que le mot risque recouvre implicitement le fait qu'une conséquence défavorable (la maladie) est précédée par une ou plusieurs conditions (facteurs de risque) qui augmentent la probabilité d'apparition de la conséquence défavorable. En première approche, on ne peut donc pas considérer comme facteur de risque tout facteur qui possède une liaison statistiquement significative avec le développement d'une maladie. Ainsi, les facteurs de risque ne doivent-ils pas correspondre à des éléments de description de la maladie elle-même mais à des variables prémonitoires devant ouvrir la voie à des possibilités de prévention (Jenicek et Cleroux, 1985). Dans cette perspective, lorsqu'une variable n'est pas modulable on parle de "marqueur de risque" et non plus de facteur de risque. En production animale intensive telle que pratiquée actuellement, on tend à soustraire les animaux de l'influence directe de l'environnement naturel et la plupart des variables environnementales deviennent potentiellement modulables car placées sous contrôle de l'Homme. Ainsi, dans ces espèces animales, les marqueurs de risque ne peuvent-ils que rarement enlever tout intérêt aux études de l'épidémiologie analytique. En regard de la recherche des facteurs de risque, les animaux de ferme élevés en claustration présentent plusieurs caractéristiques qui facilitent la conduite des opérations :

- leur environnement immédiat, outre le fait qu'il soit modulable, tend à être standardisé et relativement homogène pour tous les sujets d'un même bâtiment ou d'une même famille de bâtiments,
- l'effectif des individus que l'on peut considérer est en général important dans chacune de ces conditions standardisées,
- les cycles biologiques sont souvent considérablement plus courts que chez l'Homme : l'espérance de vie d'un porc destiné à la production de viande est de moins de 6 mois, celle d'un poulet est inférieure à 2 mois,
- enfin, les épreuves expérimentales en conditions totalement contrôlées sont de réalisation plus aisée.

Pour l'épidémiologiste en élevage intensif, la plupart des troubles de santé affectant les animaux peuvent être actuellement schématiquement rangés en deux grandes catégories.

1. Les maladies à déterminisme simple, le plus souvent monofactoriel.

Les exemples sont nombreux : maladie d'Aujeszky, peste... chez le porc, peste aviaire, maladie de Newcastle... chez les volailles, fièvre aphteuse... chez les bovins.

Ces maladies présentent en commun quelques caractéristiques générales :

- des signes cliniques et lésions souvent univoques,
- des manifestations souvent spectaculaires et d'apparition soudaine,
- l'existence de méthodes de diagnostic fiables permettant de bien repérer les sujets infectés,
- une forte contagiosité,
- une reproduction expérimentale généralement aisée.

Pour cette catégorie de maladies, la relation de causalité peut être établie sans trop de difficulté. S'agissant toutefois de maladies bien connues sous l'angle de l'étiologie, les études à entreprendre seront orientées vers d'autres cibles, la recherche de diverses circonstances permettant d'expliquer la diffusion de la maladie sur un territoire donné, la disparité des manifestations cliniques...

2. Les maladies à déterminisme complexe

Contrairement aux précédentes, elles sont communément rencontrées dans les élevages où elles sont entretenues à l'état enzootique. Leur expression peut être polymorphe et parfois récurrente. Les manifestations sont rarement généralisées et prennent une allure plus ou moins prononcée selon les individus et les élevages. C'est ainsi qu'on peut parler à leur égard de pathologies à expression progressive ou de type quantitatif. Les reproductions expérimentales sont difficiles, tout au moins sans avoir recours à des artifices expérimentaux d'inoculation très éloignés des conditions naturelles de contamination en élevage. Autant dire qu'il n'est pas fréquent qu'une étude épidémiologique débouche, en ce qui les concerne, directement sur une relation de causalité telle que définie dans la littérature médicale.

En revanche, les études pourront être orientées, comme pour les maladies de la première catégorie, vers la recherche des circonstances prioritairement associées à l'apparition et à la sévérité de l'expression des troubles ou à la propagation de ceux-ci. Cette démarche ouvre davantage sur les probabilités et la notion de risque que sur la relation de causalité. Cependant, dans le sillage d'une étude épidémiologique, une reproduction expérimentale vaut d'être tentée si les moyens matériels sont réunis. L'hypothèse causale peut alors être testée.

III. LES VOIES DE RECHERCHE DES FACTEURS DE RISQUE EN PATHOLOGIE ANIMALE

Deux hypothèses de base président à la recherche des facteurs de risque :

- les troubles de santé ne résultent pas du seul fait du hasard,
- les manifestations morbides, les caractéristiques individuelles et le milieu peuvent faire l'objet d'une description.

Ces deux hypothèses étant admises, rien ne s'oppose théoriquement aux travaux de dépistage des facteurs de risque. Il demeure néanmoins que les méthodes de recherche des facteurs de risque doivent être adaptées aux caractéristiques épidémiologiques du problème à traiter, ainsi qu'aux objectifs de chaque étude. Les commanditaires de l'étude souhaitent généralement réduire la prévalence et/ou la sévérité de la pathologie, ou tout au moins en atténuer les conséquences économiques en agissant sur les facteurs de risque appréhendés. S'agissant de la recherche de l'empreinte pathogène d'anomalies présentes chez les animaux eux-mêmes ou dans le milieu, le protocole de collecte des données devra faire l'objet d'une attention toute particulière, notamment en matière d'exhaustivité. En outre, compte tenu des enjeux, l'authenticité des facteurs de risque doit être indiscutable ; les associations mises en évidence ne doivent pas être fortuites. Enfin, les troubles de santé soumis à l'analyse sont souvent complexes, la complexité du phénomène motivant en fait fréquemment ce type de recherche. Des conséquences sont à attendre à trois niveaux.

1. La préparation du protocole

On ne peut entreprendre d'emblée une étude épidémiologique visant l'appréhension des facteurs de risque sans avoir réalisé des travaux préalables (Madec et Tillon, 1988). Ces travaux ont deux principaux objectifs :

- **L'obtention par le maître d'oeuvre d'une bonne connaissance bibliographique et pratique du problème à traiter** (acquisition d'un niveau d'expertise suffisant). Celle-ci permet d'assurer le maître d'oeuvre de la pertinence de l'entreprise projetée. Des études de cas et des travaux expérimentaux sont particulièrement recommandés. Ainsi par exemple, la transmission des troubles à des animaux à flore minimale par simple contact d'un animal venant d'un élevage atteint doit conduire, non seulement, à ne pas faire l'impasse sur la recherche des contaminants, mais encore, à orienter de manière judicieuse le protocole vers les sites de collecte de l'information.
- **La recherche d'une collection d'indicateurs épidémiologiques** qui sont autant de descripteurs de faits qu'il y a lieu par la suite de soumettre à l'investigation. La bibliographie peut déjà fournir une première liste de paramètres à considérer. Mais, bien souvent, de nouveaux champs doivent être explorés sous peine de repasser inlassablement sur les sentiers tracés par la littérature.

De nouveaux indicateurs doivent être élaborés et leur validité doit être testée (adaptation de méthodes de diagnostic à une utilisation de routine, mise au point de grilles d'évaluation...). La simplicité de mise en oeuvre et l'objectivité sont des qualités essentielles chez ces indicateurs (collecte de données "dures").

- **L'allègement du protocole** de collecte des données, tout en maintenant sa pertinence.

2. La conduite de l'enquête

La complexité des troubles conduit naturellement à envisager de manière privilégiée les **études longitudinales**, notamment les études prospectives. Celles-ci sont possibles et d'autant plus faciles chez les animaux que les cycles biologiques sont courts. Compte-tenu des exigences d'exhaustivité, elles demeurent cependant généralement assez lourdes, surtout lorsque les travaux préalables n'ont pas été suffisamment approfondis. Les **études transversales** sont néanmoins intéressantes car généralement moins laborieuses à conduire que les précédentes, ce qui permet plus facilement d'atteindre de grands effectifs d'individus. Les travaux préalables peuvent avoir permis la découverte d'indicateurs épidémiologiques descripteurs (ou "témoins") du passé récent ce qui est appréciable pour ce type d'enquête (profils sérologiques...). Les études transversales trouvent néanmoins surtout leur place lorsque le milieu présente une forte inertie sur la période considérée.

3. Le traitement des données

Les méthodes classiques de détermination des facteurs de risque peuvent être utilisées en épidémiologie animale. Mais, il arrive que pour leur application, le statisticien soit confronté à certaines difficultés.

Premièrement, la pathologie étudiée peut être décrite par des manifestations d'expression progressive : séparer les individus en deux catégories (malade/non malade ou atteint/non atteint) repose alors sur le choix d'un seuil. Ce seuil peut correspondre à la limite entre l'état normal et l'état pathologique, dans la mesure où elle peut être définie de façon stricte, mais il peut également être pertinent de distinguer les animaux sévèrement atteints des autres. Ceci peut être illustré par deux exemples étudiés chez le porc. La pathologie respiratoire est mesurée à l'échelle de l'animal par l'évaluation quantitative de la sévérité des lésions de pneumonie. Le choix d'un seuil unique est alors difficile (figure 1).

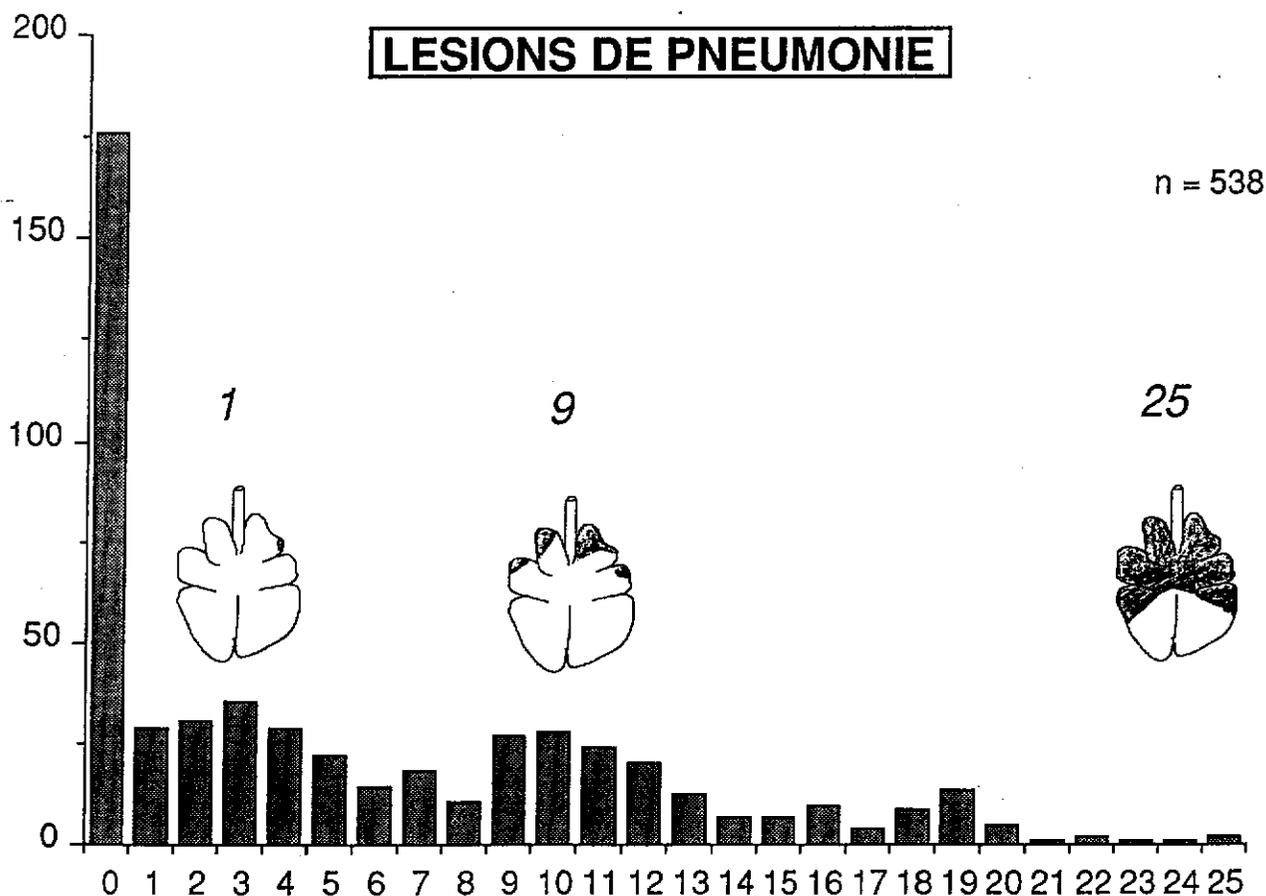
Lorsque l'individu statistique est l'élevage, la pneumonie est qualifiée, d'une part, par la sévérité des lésions observées, d'autre part, par leur fréquence. Par ailleurs, très peu d'élevages sont totalement indemnes. La seule distinction des élevages atteints et non atteints n'est donc pas pertinente.

L'exemple de la pathologie de la reproduction introduit une difficulté supplémentaire : pour évaluer les performances de reproduction à l'échelle d'un élevage, on peut mesurer en particulier le taux de retours en chaleur et la taille moyenne des portées. La limite exacte entre un état normal et un état pathologique ne peut alors pas être définie *a priori*.

Aussi, si le choix d'un seuil permet de séparer les individus statistiques en deux catégories, il est parfois délicat et entraîne une perte d'information importante.

Deuxième difficulté : en épidémiologie animale, les entités pathologiques étudiées sont souvent des syndromes définis par un ensemble de manifestations cliniques, d'origine multifactorielle. De nombreux paramètres sont alors pris en compte pour décrire les phénomènes étudiés.

Figure 1 : Exemple de l'application d'une grille de notation de 0 à 28 à une cohorte de porcs charcutiers. Dans cette population, on peut distinguer le groupe des animaux ne présentant aucune lésion (note 0) de ceux présentant des lésions : la catégorie "malade" est alors très hétérogène, elle comprend des animaux porteurs de lésions très sévères ou très discrètes. Le choix d'un autre seuil peut être déterminé par la distribution des lésions : ici on peut distinguer les animaux sévèrement atteints (note > 8) des animaux peu atteints et indemnes. Pour l'exploitation ultérieure, il apparaît souhaitable de conserver plusieurs seuils.



Ils peuvent être regroupés en deux catégories : les variables explicatives, hypothèses de facteurs de risque, qui présentent souvent des interrelations, et les variables résultantes ou variables à expliquer qui apportent une description de la maladie étudiée. Pour identifier les facteurs de risque, la recherche des liaisons entre variables prises isolément doit être complétée par la prise en considération des interrelations entre variables.

Pour prendre en compte, d'une part, le fait que les variables étudiées ne sont pas dichotomiques, d'autre part, l'existence d'interrelations, l'épidémiologiste peut s'adresser aux méthodes de la statistique inférentielle multivariée. Les méthodes de régression reposent sur la définition *a priori* d'un modèle mathématique dont on teste l'adéquation aux données. Dans cette famille, le choix d'une méthode dépend de la nature des variables à traiter (encadré 1). Elles sont complétées par les méthodes de l'analyse de variance. Les analyses permettent d'identifier les facteurs de risque, de leur affecter un coefficient pour quantifier l'intensité de la relation, et de déterminer un seuil de signification des résultats. Toutefois, elles présentent également des limites qu'il convient de préciser. La complexité des modèles mathématiques ne permet pas actuellement, lorsque l'on prend en compte les interrelations entre variables, d'étudier simultanément un nombre important de variables.

Encadré 1 : Les méthodes inférentielles multivariées.

PRINCIPE : 1. Définir un modèle mathématique
 2. Tester son adéquation aux données lors de simplifications

METHODES ET UTILISATION	Variables explicatives	Variables résultantes
Régression linéaire multiple	n quantitatives	1 quantitative
Régression log-linéaire	n qualitatives	1 qualitative
Régression logistique	n qualitatives ou quantitatives	1 qualitative binaire
Analyse de variance	n qualitatives	1 quantitative

RESULTAT :

FACTEURS DE RISQUE : variables qui restent dans le modèle simplifié, caractérisées par un coefficient.

Ainsi doit-on se limiter le plus souvent à 4 ou 5 variables. Par ailleurs, il n'est pas possible d'introduire dans une même analyse plusieurs variables à expliquer ; cela impose de décrire la pathologie étudiée par une seule variable. Dans le cas de régression logistique, il est, de plus, nécessaire qu'elle soit binaire (présence/absence). D'autre part, au cours des différentes étapes de la modélisation, le statisticien doit choisir entre plusieurs options dont dépendent les résultats ; par exemple, il doit, soit privilégier l'analyse des interactions, soit privilégier la spécificité des mises en classe.

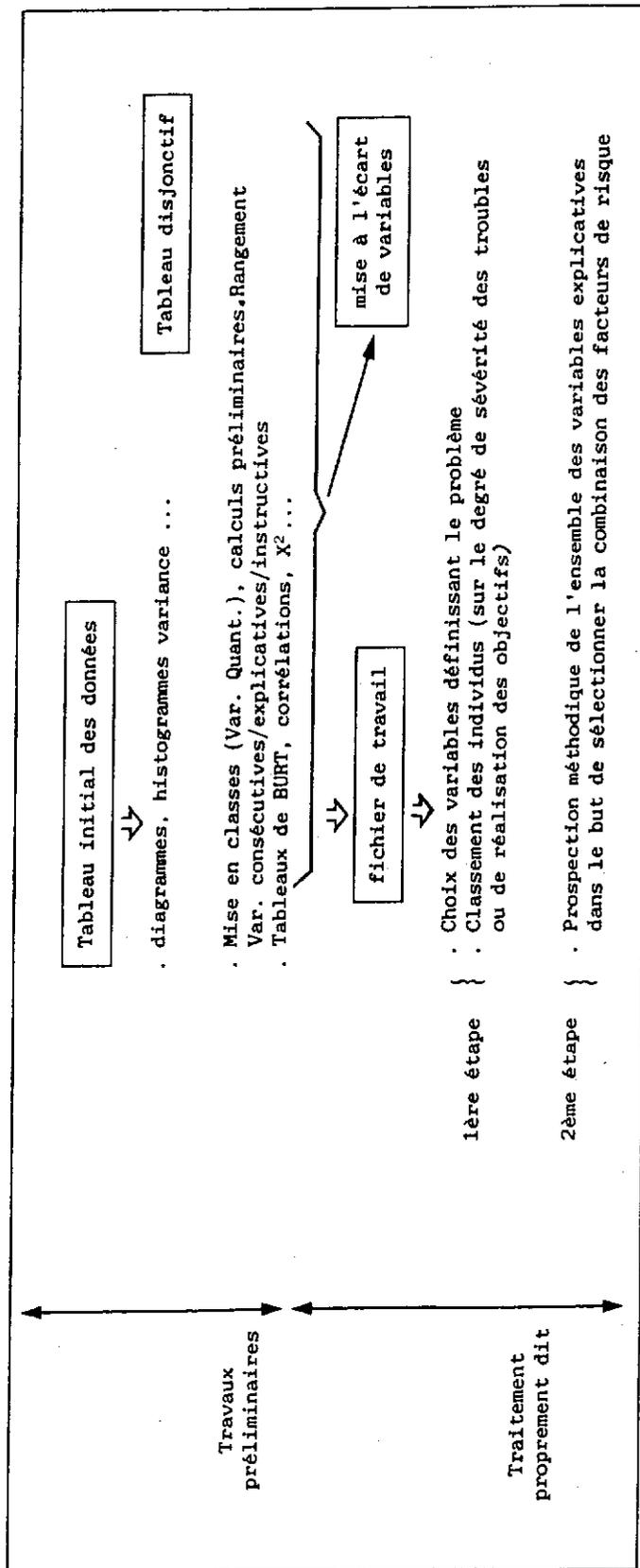
L'analyse de la sensibilité des modèles à ces options met alors parfois en évidence leur manque de robustesse (Enevoldsen et coll., 1990). Enfin, pour la mise en oeuvre et l'interprétation des méthodes inférentielles, il faut distinguer les variables qui sont des facteurs de risque potentiels, des facteurs de confusion, paramètres associés au phénomène étudié, mais dont le rôle dans l'induction d'un risque est écarté. Or, les exemples fournis par les études d'épidémiologie animale montrent, d'une part, la difficulté à écarter *a priori* le rôle d'un facteur et, d'autre part, l'intérêt d'identifier les indicateurs de risque, associés à la pathologie, mais dépourvus de rôle causal.

Aussi, en complément des méthodes inférentielles, le recours à d'autres méthodes statistiques a-t-il été envisagé : les méthodes descriptives multivariées ou méthodes de l'école française d'analyse des données.

IV. UTILISATION DES METHODES DESCRIPTIVES MULTIVARIEES

Le schéma général du processus de recherche des facteurs de risque est mentionné à la figure 2. Les méthodes multivariées n'interviennent qu'après une phase de travaux préliminaires.

Figure 2 : Schéma général des opérations aboutissant à la sélection des facteurs de risque.



1. Les travaux statistiques préliminaires

Ils sont en général assez longs mais revêtent un caractère de nécessité. Ils requièrent la participation active des différents partenaires de l'équipe d'épidémiologie assurant la maîtrise d'oeuvre car des options importantes sont parfois à prendre, engageant la responsabilité du maître d'oeuvre.

- Les données sont rangées selon leur nature en regard du phénomène pathologique. Ainsi va-t-on **répartir les variables enregistrées en trois groupes** : les variables consécutives, les variables explicatives et les variables instructives.
 - . Les **variables consécutives** sont des paramètres qui décrivent la maladie ou le phénomène qu'on se propose d'étudier, ou qui sont en aval de son apparition. Ainsi, en général les variables consécutives se recrutent de manière privilégiée parmi les symptômes, les lésions et les performances zootechniques.
 - . En revanche, dans le tableau des données, il existe une multitude d'éléments susceptibles de précéder ou d'accompagner, seuls ou associés, la mise en route ou le développement des manifestations pathologiques. Lorsqu'en aucun cas, ils ne peuvent en être des conséquences (pour autant que la vérification soit possible), ces variables sont dites **explicatives**. Il s'agit principalement des caractéristiques des individus avant l'apparition des premiers signes (filiation, poids, sérologie...) ainsi que celles de son environnement avant et au cours de la phase d'expression perceptible de la maladie (alimentation, bioclimat...).
 - . Enfin, il existe une autre catégorie de variables qui peuvent être à la fois "explicatives" et "consécutives". Ces variables, généralement peu nombreuses lorsque la chronologie des faits observables est convenablement établie au cours des travaux préalables, et le protocole adapté en conséquence, sont dites "**variables instructives**".

Les premières **descriptions statistiques** sont entreprises au cours de cette phase. Les variables sont envisagées isolément (paramètres de dispersion) puis deux à deux (tableaux de comptage, tests divers : χ^2 , corrélations...). Des calculs de nouvelles variables sont également réalisés à partir des données élémentaires collectées (calculs de délais à partir de dates ; d'index pré-établis, à partir de données de base...). Ces travaux permettent une première connaissance du contenu du tableau de données et ils aboutissent à confectionner un "fichier de travail", de dimension plus réduite puisque débarrassé des colonnes nulles ou constantes ainsi que de toutes les variables sans intérêt pour la suite des travaux. Les variables quantitatives sont éclatées en classes et on aboutit à un tableau disjonctif complet.

2. Le traitement des données proprement dit

Jusque là ce sont surtout l'**analyse factorielle des correspondances (AFC)**, l'**analyse des correspondances multiples (ACM)** et la **classification ascendante hiérarchique (CAH)** qui ont été utilisées par notre équipe (Benzecri, 1976 ; Jambu et Lebeaux, 1983 ; Lebart et coll., 1985).

L'analyse comporte deux étapes :

La **définition des "objectifs"** en regard de la maladie. En d'autres termes, on veut trouver le paramètre ou le groupe de paramètres qui permet au mieux de décrire la maladie.

Plusieurs variables "consécutives" sont susceptibles de participer à la formulation des objectifs. Celles-ci doivent s'avérer pertinentes et suffisamment liées pour pouvoir être étudiées simultanément. Dans le cas de syndromes particulièrement complexes, il peut arriver en effet que le spécialiste, à la lecture des premiers résultats, s'aperçoive de sérieuses divergences dans le déploiement des variables qu'il voulait initialement considérer simultanément comme cible. Il doit alors, au besoin, étudier séparément les groupes d'affinités de variables consécutives.

Quoi qu'il en soit, à l'issue de cette étape, les individus (élevages, animaux, lots d'animaux...) seront rangés selon la sévérité de la pathologie ou certaines caractéristiques de celle-ci.

La seconde étape consiste en une **prospection méthodique du groupe des variables explicatives**, dans le but d'en extraire la **combinaison** de variables qui permet au mieux de discriminer les familles d'individus obtenues à la première étape. Une bonne connaissance du dossier, à partir de l'examen détaillé des corrélations et des tableaux de comptage, permet l'établissement d'une carte de départ composée de quelques variables à bonne contribution. Cette contribution est mesurée par la comparaison de leur comportement dans l'analyse avec celui des variables consécutives retenues. Avec la mise en place d'un "fond de carte", le comportement de nouvelles variables est étudié. Certaines, dont l'influence est favorable sur la distribution des points, sont conservées, les autres sont provisoirement écartées. La base de travail reste un ensemble mouvant au point que certaines variables initialement retenues doivent être abandonnées au bénéfice d'une ou de plusieurs autres et inversement. La retenue ou l'abandon des variables sont fondés sur la distribution des observations induite par cette incorporation ou par ce rejet.

La décision est basée sur une comparaison avec la répartition obtenue sur les variables consécutives à l'issue de la première étape. La comparaison est en effet possible puisque chaque observation conserve son identité. En outre, la constitution à partir des observations de groupes d'affinités permet l'attribution d'une symbolique pour chacun d'eux. Plus les groupes "éloignés" sur la carte des objectifs sont de nouveau éloignés l'un de l'autre sur la carte des variables explicatives et plus le chemin parcouru dans la recherche des facteurs de risque est important. Les groupes d'individus peuvent également être avantageusement visualisés par des ellipses de garde.

La séquence des analyses peut être longue, mais on peut limiter raisonnablement leur nombre :

- En approfondissant les travaux de description préalable.
- En mettant en place une organisation rigoureuse. La tenue d'un véritable plan de route est nécessaire afin de récapituler clairement les combinaisons de variables testées et les résultats obtenus. Ces données vont servir à l'élaboration des nouveaux projets d'incorporation.
- En ayant recours à une technique de localisation sur le graphe de l'analyse de variables dites "supplémentaires" (Fenelon, 1981). Ces dernières ne participent aucunement à la déformation du nuage mais leur situation sur la carte est cependant fonction des "relations" entretenues avec les variables analysées. Cette technique permet la présélection de variables explicatives en vue des analyses projetées. De plus, les variables consécutives sont ainsi introduites ce qui constitue une autre manière de visualiser le progrès réalisé au fil des analyses par la simple mesure de leur répartition sur le graphe.

- En procédant à une étude de détail de chacune des analyses (contributions absolues et relatives des variables, étude des groupes d'affinités sur les différents facteurs...). Une ou plusieurs classifications hiérarchiques sont couplées à chaque analyse des correspondances pour aider à leur interprétation. Le comportement des variables explicatives peut ainsi être appréhendé sur les axes factoriels sur lesquels contribuent en priorité les variables consécutives (incorporées en qualité de variables supplémentaires).

La séquence est interrompue lorsque l'introduction ou l'abandon d'une variable ne permet plus de se rapprocher sensiblement du classement des observations obtenu au terme de la première étape. La combinaison ultime des variables explicatives est appelée combinaison des facteurs de risque.

3. Signification des facteurs de risque ainsi obtenus

La signification qu'on peut donner aux facteurs de risque ne peut être considérée qu'en regard de la méthode d'obtention de ces facteurs. En ce qui concerne l'approche décrite ci-dessus, on peut faire les remarques suivantes :

- Au cours de l'enquête épidémiologique, les données sont enregistrées *in situ* dans les élevages. Les phénomènes pathologiques surviennent spontanément, sans aucun artifice, et indépendamment de la volonté du maître d'oeuvre. En cela, ces études sont à distinguer de l'expérimentation basée sur un plan expérimental (pour lequel la nature, la dose de l'inoculum, le moment de l'inoculation sont fixés à l'avance...). L'objet du traitement statistique est ici de faire apparaître au terme de l'étude comparative, la meilleure combinaison des paramètres appartenant au groupe des variables explicatives et qui sont associés au phénomène pathologique. Ces paramètres doivent en toute logique entretenir avec le problème considéré des relations statistiques privilégiées. Les méthodes mathématiques qui ont été utilisées permettent de mettre en évidence ces "liaisons" ou plus généralement les "structures" qui existent au sein d'un tableau de données. Le traitement ne concerne ici nullement la recherche du taux d'explication de la variation totale par tel ou tel facteur, ni la recherche d'une signification statistique comme c'est le cas dans la plupart des essais.
- Les techniques de calcul utilisées, qui font largement appel aux méthodes multidimensionnelles, conduisent à la notion de "profil" d'individu, défini par les valeurs qu'il prend pour plusieurs variables. Ainsi doit-on parler de "profil à risque", la probabilité de voir se développer le problème chez un individu étant fonction des valeurs obtenues sur l'ensemble des paramètres considérés simultanément. En conséquence, en vue d'une application pratique, **il n'est pas souhaitable d'envisager un "facteur de risque" indépendamment des autres.**
- La démarche rigoureuse de préparation, de réalisation puis de sélection des facteurs de risque vise à conférer à ces derniers une valeur prédictive, **l'objectif final étant de bâtir un canevas de travail utilisable dans la lutte contre les maladies.** Pour cela, il importe au maître d'oeuvre de posséder une bonne connaissance de la maladie et notamment de la chronologie des faits observables, car il y a lieu d'éviter l'aboutissement à la sélection de "facteurs de risque" qui ne sont en réalité que des manifestations du processus morbide. En réalisant une confusion entre variables consécutives et variables explicatives, on peut arriver en effet à "expliquer" une maladie par ses propres symptômes. Imaginons par exemple que les quintes de toux chez le porc à l'engrais soient étroitement associées aux lésions de pneumonie ; s'agit-il d'un facteur de risque ? La fréquence des toux peut certes posséder une bonne valeur prédictive à l'égard de la pneumonie décelée à l'abattoir mais dans la perspective d'une stratégie de lutte, cette découverte ne présente pas grand intérêt.

Il ne peut s'agir d'un facteur de risque au sens où nous l'entendons puisque combattre la pneumonie revient en fait à éviter l'apparition de symptômes de cette maladie.

- A l'issue de l'obtention de la combinaison de facteurs de risque, celle-ci est validée par une mise à l'épreuve sur un autre échantillon d'élevages. L'étape de la validation des résultats vise à accumuler les preuves du caractère réel et non fortuit de l'association. La démonstration est parfois éloquent, la correction des facteurs de risque permettant aux signes cliniques de s'estomper.
- Enfin, les facteurs de risque n'ont ni valeur définitive ni valeur universelle. Il y a lieu d'envisager des actualisations périodiques ainsi que des vérifications de pertinence dans les zones géographiques autres que celle qui a servi de support aux études initiales.

CONCLUSION

La recherche des facteurs de risque représente une activité essentielle des épidémiologistes, car elle doit déboucher sur des moyens de prévention des troubles de santé. Les travaux réalisés sur animaux, notamment en élevage confiné intensif, sont probablement plus faciles à conduire que ceux entrepris chez l'Homme, en raison de la brièveté des cycles et de plus larges possibilités pour objectiver les phénomènes et intervenir. Différentes voies sont possibles pour la mise en évidence des facteurs de risque. Celles-ci concernent tant le protocole d'étude que la méthode de traitement statistique des données.

Les méthodes conventionnelles trouvent leur place dans l'étude de phénomènes simples, globalement élucidés mais pour lesquels il y a lieu de préciser l'impact relatif d'un nombre restreint de variables. En revanche, pour l'étude de troubles complexes, il peut être judicieux de mettre en oeuvre à la fois les méthodes inférentielles et les méthodes descriptives multivariées dans le processus du traitement des données.

BIBLIOGRAPHIE

- BENZECRI J.P. (1976).- L'analyse des données.
Tome I : La taxinomie. Dunod éd., Paris
Tome II : L'analyse des correspondances. Dunod éd., Paris.
- ENEVOLDSEN C., GROHN Y.T. and THYSEN I. (1990).- Proceedings of the SVEPM Conference, Belfast, p. 133-144.
- EVANS A.S. (1978).- Causation and disease : A chronological journey. Am. J. Epidemiol., **108**, 249-258.
- FENELON P. (1981).- Qu'est-ce que l'analyse des données ? Lefonen éd., Paris.
- GRUNDY P.F. (1973).- A rational approach of the "at risk" concept. The Lancet Déc. 29, 1489.

- JAMBU M. and LEBEAUX N.D. (1983).- Data analysis and cluster analysis. North Holland ed, Amsterdam.
 - JENICEK M. et CLEROUX R. (1985).- Epidémiologie clinique, clinimétrie. Maloine éd., Paris.
 - LEBART L., MORINEAU A. and WARWICK L. (1985).- Multidimensional descriptive data analysis. John Wiley ed., New-York.
 - MADEC F. et TILLON J.P. (1988).- Ecopathologie et facteurs de risque en médecine vétérinaire. Analyse rétrospective (1977-1987) de l'expérience acquise en élevage porcin intensif. Rec. Méd. Vét., **164**, 607-616.
-