

# L'ÉCHANTILLONNAGE EN ENQUÊTE ANALYTIQUE

Marie-Liesse Lassauzet <sup>[1]</sup>, C. Lopez <sup>[2]</sup>,  
B. Faye <sup>[3]</sup> et F. Madec <sup>[4]</sup>

## Résumé

*Les méthodes d'échantillonnage utilisées dans les enquêtes analytiques sont passées en revue, sur les aspects des plans de sondage et des tailles d'échantillons, selon le type d'étude, prospective ou rétrospective, et selon l'objectif de l'enquête, descriptif ou explicatif.*

*Des procédures de validation sont recommandées pour pallier le non respect, en pratique, des méthodes d'échantillonnage, constaté dans beaucoup d'études en médecine vétérinaire.*

*Des exemples d'études réalisées en France sont présentés.*

## Summary

*Sampling methods used in analytical epidemiological studies are reviewed with respect to sampling schemes and samples size, according to the type of study, prospective or retrospective, and to the study objective, exploratory or explanatory. Since, in veterinary medicine, most analytical studies present deficiencies in their sampling methods for practical reasons, some validation procedures are recommended. Examples from studies performed in France are presented.*



L'échantillonnage est une étape essentielle des enquêtes épidémiologiques, rendue nécessaire lorsqu'il n'est pas possible de recenser et de suivre l'intégralité de la population-cible.

Il consiste à sélectionner par diverses méthodes des "unités d'échantillonnage" (animaux, lots, troupeaux, ...) sur lesquelles sera réalisée l'étude, à défaut de l'être sur la totalité de la population. L'enquête est alors menée sur l'ensemble des unités échantillonnées ("l'échantillon"), avec l'intention implicite d'extrapoler les résultats observés à la population parente.

Ce passage de l'échantillon à la population ("l'inférence statistique") présuppose que l'échantillon doit "ressembler" à la population d'origine. Il doit être, en quelque sorte, une population miniature.

Cette condition est essentielle en épidémiologie descriptive lorsqu'il s'agit d'estimer des prévalences ainsi que leur précision. Elle l'est aussi en épidémiologie analytique lorsque l'on veut mettre en évidence des relations entre facteurs et pathologie et quantifier l'importance des facteurs de risque.

[1] CIBA-GEIGY, Santé Animale, AH 3.4, 4002 Bâle, Suisse

[2] Institut de l'Élevage, Service Biométrie, 149 rue de Bercy, 75595 Paris cedex 12, France

[3] Laboratoire d'Écopathologie, INRA, Centre de Recherches de Clermont-Theix, 63122 Saint Genès Champanelle, France

[4] CNEVA, Station de Pathologie Porcine, B.P. 53, 22440 Ploufragan, France

Tout protocole d'échantillonnage doit définir de façon précise les modalités de sélection des unités mais aussi leur nombre. Le contrôle *a priori* de la taille des échantillons est en effet la condition nécessaire permettant de maîtriser de la puissance des tests dans les analyses ultérieures. Le coût de recueil des observations peut être aussi un élément limitant important dans la fixation des tailles d'échantillons.

Meydrech et coll. [53] et Pike et coll. [56] proposent des procédures permettant de prendre en compte, dans le calcul des tailles, des coûts différents suivant l'exposition ou non au facteur étudié et suivant que l'unité statistique est considérée comme atteinte par la maladie ou pas.

Le protocole doit enfin prévoir une phase de validation *a posteriori* permettant de contrôler la qualité des échantillons obtenus.

Dans une première partie, nous allons présenter les différentes méthodes d'échantillonnage pratiquées en épidémiologie analytique ainsi que le calcul des tailles d'échantillons suivant le type d'enquête et les techniques actuelles de validation des échantillons. Dans une seconde partie, les expériences en la matière du Laboratoire d'écopathologie de l'INRA de Theix et de la Station de pathologie porcine de Ploufragan permettront d'illustrer la pratique de ces méthodes.

## PREMIERE PARTIE : Les méthodes d'échantillonnage en épidémiologie analytique

### I - ECHANTILLONNAGE ALEATOIRE ET NON ALEATOIRE

La notion intuitive de ressemblance de l'échantillon avec la population parente évoquée précédemment est définie de façon rigoureuse, dans la théorie de l'échantillonnage, par la notion de "représentativité". Cette notion assure que **tout individu** de la population peut être **potentiellement** un élément de l'échantillon. Cette condition est remplie par le tirage au sort des individus. La théorie des sondages offre toute une panoplie de modes de tirages au sort : le tirage aléatoire simple, systématique, stratifié, à plusieurs degrés, en grappes ... . Il appartient aux responsables de l'enquête de choisir le plan de sondage adéquat en fonction des objectifs de l'étude mais aussi des coûts de recueil des observations.

La lourdeur de certaines études nécessite toutefois un degré de participation des éleveurs qui rend quasi impossible, en pratique, le tirage au sort des élevages ou des animaux. La non

disponibilité d'informations sur les niveaux d'exposition et/ou sur la pathologie peut également rendre inapplicable la constitution d'échantillons aléatoires. Des méthodes "non aléatoires" sont alors proposées basées sur des "choix raisonnés" des unités de la population. Ainsi, des élevages peuvent être choisis parce qu'ayant déjà participé à des études antérieures et que l'information y est facile d'accès et relativement fiabilisée, ce sont les "*convenience sampling*" (Martin, [49]). Ou bien, dans les enquêtes analytiques, le choix des unités est basé sur la connaissance soit de l'exposition aux facteurs soit de la pathologie objet de l'étude, ce sont les "*purposive sampling*" (Martin, [49]).

Ces techniques sont parfois les seules applicables en pratique. Elles présentent l'inconvénient de ne pas garantir l'absence de biais de sélection et donc nécessitent une validation *a posteriori* des échantillons très soignée.

## II - LES TYPES D'ENQUÊTES ANALYTIQUES DE BASE

On distingue classiquement trois types d'enquêtes analytiques (Fleiss [30]) :

- les enquêtes transversales ("*cross sectional studies*"),
- les enquêtes prospectives ou études de cohorte ("*cohort studies*"),
- les enquêtes rétrospectives ou études cas/témoin ("*case-control studies*").

Ces trois classes génériques d'enquêtes se distinguent principalement par le mode de sélection des unités statistiques.

♦ Dans les enquêtes transversales seul le nombre total d'individus est contrôlé *a priori*. Ceux-ci sont en théorie tirés au hasard dans la population. Les distributions suivant les facteurs d'exposition et la maladie étudiée sont aléatoires.

L'objectif est ici essentiellement descriptif avec une possible étude de l'association entre facteurs d'exposition et la pathologie (étude de prévalence) mais sans aucune maîtrise de la puissance des tests correspondants, conséquence du non-contrôle de la distribution des facteurs. Cet objectif est généralement qualifié "d'exploratoire".

Ce type d'enquête est particulièrement peu pertinent lorsque l'on étudie des maladies rares. Il concerne plutôt, en épidémiologie analytique, les situations où l'on ne connaît pratiquement rien des facteurs de risque de la pathologie. Il ne sera pas développé dans la suite.

♦ Dans les enquêtes prospectives l'on contrôle généralement le nombre d'individus exposés à un facteur susceptible d'être facteur de risque de la pathologie étudiée. L'échantillonnage est aléatoire stratifié, chez les individus exposés et non exposés, ou à choix raisonné (enquêtes à objectif explicatif ou expérimentation (essais cliniques)). Il est aléatoire ("*Lot Quality Assurance Sampling*") pour les enquêtes de surveillance (Reinke, [58]).

♦ Dans les enquêtes rétrospectives la distribution de la pathologie est contrôlée. L'échantillonnage est aléatoire chez les malades et les témoins ou à choix raisonné, avec ou sans appariement.

♦ Enfin, citons les études "prospectives-rétrospectives" où les cas sont appariés à des témoins provenant d'une cohorte obtenue par tirage aléatoire dans la population de référence.

Ces différentes études sont récapitulées dans le Tableau I. Elles vont être détaillées dans la suite.

Tableau I : Les différents types d'enquêtes et le mode de sélection des unités d'échantillonnage

Types d'enquêtes	Etudes	Objectif	Méthode de sélection
Transversale		exploratoire	Aléatoire (simple, stratifié, systématique, en grappes et combinaisons).
Prospective		exploratoire	Aléatoire (simple, stratifié, systématique, en grappes et combinaisons).
	exposé/non exposé	explicatif	Choix raisonné
	essai clinique essai d'intervention	explicatif	Choix raisonné
	surveillance	exploratoire	Aléatoire (Lot quality Assurance sampling)
Rétrospective	cas/témoin	exploratoire/explicatif	Aléatoire (avec ou sans appariement)
Prospective/Rétrospective		explicatif	Aléatoire (avec appariement)

### III - METHODE DE SELECTION ET TAILLE DE L'ECHANTILLON

#### A - ETUDES PROSPECTIVES A OBJECTIF EXPLORATOIRE

Ce sont des études destinées à mesurer la fréquence de diverses maladies dans des élevages et qui veulent associer ces maladies à des facteurs de risque démographiques (eg. : âge, race, sexe), de production/reproduction, de management (eg. : habitat, nourriture, type de salle de traite) ou climatiques. En général, ces études se sont déroulées sur une période de 1 à 5 ans. Au début de l'étude, un questionnaire est rempli et on indique à l'éleveur comment recueillir les données. Puis des visites sont faites tous les 1 à 3 mois pour vérifier et contrôler les données. Peu ou pas d'échantillons biologiques sont recueillis durant l'enquête en dehors de ceux prélevés normalement (eg. : contrôle laitier, prophylaxies).

L'échantillonnage de ces études devrait être aléatoire puisque le but pragmatique de ces études est de prendre des décisions sur les actions à favoriser, dans une région ou un pays, pour améliorer l'état de santé des animaux. C'est le cas pour la plupart des enquêtes faites au Canada et aux Etats-Unis ; par exemple Meek et coll. [52] dont les résultats ont été ultérieurement aussi utilisés par Bigras-Poulin et coll. [10], Mc Dermott et coll. [50] [51] ; Gardner et coll. [32] dont les résultats ont aussi été utilisés par Sischo et coll. [65], Hird et coll. [36], Kaneene et Hurd [39], [40], [37].

Dans ces études, il n'est pas toujours clairement mentionné le biais possible de la base de sondage par rapport à la population de référence, le pourcentage de troupeaux sélectionnés qui ont refusé de participer ainsi que les modalités de sélection d'autres troupeaux après un refus. Mc Dermott et coll. [50] ont eu, sur 300 troupeaux sélectionnés, un taux d'acceptation de 70 % et n'ont pas remplacé les troupeaux des éleveurs qui avaient refusé. Meek et coll. [52] ont sélectionné 110 troupeaux et ont obtenu un taux d'acceptation de 75 %. Pour obtenir une taille d'échantillon de 110 troupeaux, les troupeaux de remplacement ont été sélectionnés de façon aléatoire dans la même liste de départ mais il n'est pas fait mention du succès des échantillonnages suivants. Sur 5 ans, 10

troupeaux ont arrêté le suivi, 5 pour vente ou perte du troupeau et 5 par refus de continuer (Meek et coll., [52]). Les études du 'National Animal Health Monitoring System' ont eu des taux d'acceptation de 53 et 47 % en Californie (Gardner et coll., [32]), (Hird et coll., [36]) et de 89,2 % au Michigan (Kaneene et coll., [39]). En Californie, les troupeaux les plus proches de ceux pour lesquels les éleveurs avaient refusé leur participation (et de même catégorie de taille) ont été sélectionnés. Le succès de ce mode de remplacement n'est pas mentionné. Au Michigan, il n'y a pas eu de nouvelle sélection.

Les méthodes d'échantillonnage dans les études mentionnées ci-dessus étaient de type : aléatoire simple (Hurd et coll., [37]), (Meek et coll., [52]), stratifiée sur la région (Mc Dermott et coll., [50]) ou stratifiée sur la taille des troupeaux avec sondage à taux constant (Gardner et coll., [32]). Le type d'échantillonnage affecte l'estimation des variables que l'on veut évaluer. La stratification sur la taille du troupeau, notamment, permet d'améliorer l'estimation des variables liées à la taille du troupeau (Akhtar et coll., [1]).

La taille de l'échantillon est en général justifiée par le personnel disponible en tenant compte parfois des valeurs hypothétiques des paramètres à estimer et de leur précision (Gardner et coll., [32]), (Akhtar et coll., [1]). Après avoir sélectionné la taille de leur échantillon, Mc Dermott et coll. [50] discutent son influence sur la précision des paramètres qu'ils veulent estimer et sur la puissance des tests statistiques qu'ils veulent effectuer. En épidémiologie humaine, des chercheurs travaillent sur la détermination de la taille de l'échantillon pour les études multivariées, surtout celles visant à évaluer le risque relatif en fonction de différents facteurs de risque (Greenland et coll., [34]), (Lemeshow et coll., [44]), (Lubin et coll., [46]).

Dans d'autres études, les troupeaux ont été sélectionnés par choix raisonné. C'est le cas, par exemple, d'études faites en France ou en Suède (Barnouin, [6]), (Pluvinage et coll. [57]), (Philippot et coll. [55]), (Bendixen et coll., [9]), (Andersson et coll., [3]), ou au Canada (Dohoo et coll., [18], [19], [20], [21], [22]). La taille de l'échantillon n'est pas justifiée. La taille de l'échantillon, entre 900 et 1 000 troupeaux

(Bendixen et coll., [9]), ou le nombre de tests biologiques effectués (Andersson et coll., [3]), (Dohoo et coll., [18]) peuvent expliquer l'absence d'échantillonnage aléatoire. La liste des études dans cette catégorie est loin d'être complète, beaucoup d'études épidémiologiques sont dans ce cas.

Dans toutes les études, l'ensemble des animaux des troupeaux sélectionnés, éligibles pour l'étude, ont été suivis.

Il faut noter que ce type d'étude ne permet pas de quantifier et de tester de façon optimale les effets des facteurs de risque sur l'apparition de la maladie.

### B - ETUDE PROSPECTIVE EXPOSES/NON EXPOSES

Ce sont des études sur un ou quelques troupeaux sélectionnés par choix raisonné où le suivi est intensif et inclut des facteurs de risque biologiques précis, des examens cliniques ou nécropsiques approfondis. Les données récoltées sont donc d'excellente qualité, le contexte est quasi-expérimental et les relations de cause à effet peuvent être établies. Cependant, la généralisation n'est pas directement possible et nécessite des procédures de validation (Barnouin [7]).

### C - ESSAIS CLINIQUES ET ESSAIS D'INTERVENTION

Ces expérimentations sont des études prospectives ayant pour objectif de prouver l'efficacité d'un médicament ou d'un vaccin. Elles diffèrent des études prospectives à objectif explicatif dans le sens où les facteurs de risque autres que le traitement sont généralement connus, leurs effets sont contrôlés au niveau de l'allocation du traitement et pris en compte dans l'analyse statistique.

La sélection des troupeaux ou des cliniques vétérinaires dans lesquels sont pratiqués ces essais n'est pas aléatoire car ces essais doivent être réalisés selon des Bonnes Pratiques Cliniques strictes ; par contre, l'affectation du traitement est aléatoire. En général, tous les animaux éligibles d'un troupeau ou d'une clinique sont

inclus jusqu'à la date de fin d'essai ou l'obtention d'un nombre suffisant d'animaux.

Si l'on compare le pourcentage de cas observés dans le groupe traité et dans le groupe témoin, des formules indiquent la taille de l'échantillon selon les pourcentages attendus dans le groupe traité et dans le groupe témoin et les seuils de risque de première et de deuxième espèce. Des méthodes pour calculer la taille de l'échantillon ont aussi été développées lorsque l'on compare les survies (Schoenfeld, [62]), (Lakatos et coll., [43]), (Yateman et coll., [70]), ainsi que lorsque l'on veut évaluer le risque relatif (Blackwelder, [11]) pour des dispositifs expérimentaux plus complexes (eg. : cross-over, relations doses/effets, plusieurs traitements) (Bauer, [8]), (Day et coll., [17]), (Royston, [59]). Une revue des différences méthodes a été effectuée par Donner [23]. Enfin, lorsque l'on désire évaluer, par exemple, l'efficacité d'un vaccin pour les animaux d'un troupeau, la corrélation entre les animaux intra troupeau est prise en compte par Donner [24] et Shoukri [64] dans le calcul des tailles d'échantillons.

Les essais d'intervention, ou essais de communauté, sont des études prospectives ayant pour objectif de prouver l'efficacité d'une intervention de prévention, non-thérapeutique, sur une population. Comparés aux essais cliniques, les essais d'intervention se déroulent dans des conditions peu ou pas contrôlées, avec des groupes de comparaison qui diffèrent souvent sur plusieurs facteurs autres que celui dont on veut tester l'effet et qui sont de petite taille, car l'unité statistique est une communauté ou un troupeau. La méthodologie d'échantillonnage est discutée par Donner et coll. [24] et la taille de l'échantillon par Koepsell et coll. [41].

### D - ETUDE PROSPECTIVE DE SURVEILLANCE

Ces études sont destinées à déterminer le taux d'une maladie ou le degré d'immunisation pour pouvoir identifier des sous-groupes de la population qui ont besoin de recevoir une attention médicale particulière. La méthode d'échantillonnage est aléatoire mais diffère des méthodes traditionnelles utilisées pour l'estimation de prévalence [Reinke, 58].

### **E - ETUDE RETROSPECTIVE CAS/TEMOINS**

Ceux sont les études cas/témoins. Souvent, tous les cas sont sélectionnés et les techniques d'échantillonnage pour les témoins peuvent être simples, stratifiées ou, en grappes (Graubard et coll., [33]). Les tailles d'échantillon sont calculées comme pour les études prospectives (Greenland, [34]), (Lemeshow et coll., [44]), (Lubin et coll., [46]). Si les témoins sont sélectionnés par des techniques d'appariement, l'échantillonnage ne doit pas être systématique (Weinberg et coll., [69]). Les tailles d'échantillon sont calculées en tenant compte de cet appariement (Weinberg et coll., [69]), (Smith et coll., [66]).

### **F - ETUDE PROSPECTIVE/RETROSPECTIVE A OBJECTIF EXPLICATIF**

Ceux sont des études prospectives où des facteurs de risque, faciles à mesurer, sont enregistrés sur l'ensemble de la cohorte, sélectionnée de façon aléatoire, dans un premier temps. Ensuite, lorsque des cas surviennent, ils sont appariés à des témoins au sein de la cohorte et des facteurs plus précis sont alors mesurés sur ces cas/témoins (Flanders, [29]). Cette méthode de sélection permet d'avoir moins de biais que dans le cas d'une étude cas/témoin normale et augmente la faisabilité de l'étude en n'observant qu'une partie de la cohorte.

## **IV - VALIDATION DE LA QUALITE DE L'ECHANTILLON**

En général, il est possible d'identifier des faiblesses dans toutes les études analytiques. Celles-ci peuvent se situer au niveau du mode de sélection de l'échantillon, de sa taille, de la qualité des données collectées, de la méthode statistique utilisée et de l'interprétation des résultats. C'est pourquoi il est important de valider les résultats. Cependant, si les résultats diffèrent, il n'est pas toujours possible d'en connaître la cause exacte.

Il y a cinq possibilités, non exclusives, de valider la qualité de l'échantillon.

### **A - COMPARAISON AUX ESTIMATIONS D'UNE REGION, D'UN PAYS**

Pour les échantillons obtenus par choix raisonné, les estimations de paramètres comme la production laitière, la taille de troupeaux, l'intervalle vêlage-vêlage, sont comparées à des estimations connues au niveau régional ou national pour vérifier la représentativité des échantillons. C'est une méthode simple et peu coûteuse. Une comparaison positive n'assure pas cependant une représentativité au sens strict, elle permet simplement d'être assuré d'une absence de biais sur les paramètres les plus importants (Barnouin et Brochart, [4]).

### **B - BOOTSTRAP, JACKKNIFE ET AUTRES METHODES DE RE-ECHANTILLONNAGES**

Ces méthodes, développées par Efron [26], sont encore peu utilisées en épidémiologie. De façon schématique, des échantillons sont tirés de façon aléatoire à partir de l'échantillon initial, généralement grand, et l'analyse statistique est faite sur chaque sous-échantillon. Les facteurs sélectionnés dans la majorité des analyses peuvent être considérés comme importants (Schwartzbaum et coll. [63]). Ces méthodes ne peuvent toutefois pas éliminer un biais déjà existant dans la sélection de l'échantillon initial. Elles permettent toutefois de tester la stabilité des résultats.

### **C - VALIDATION SUR DE NOUVEAUX ANIMAUX**

Les études analytiques permettent d'identifier des facteurs de risque importants par rapport à une maladie précise. Ces résultats doivent permettre de mettre en place un plan de contrôle sur un grand nombre d'animaux et de quantifier l'effet de ce contrôle sur la maladie. Si les résultats obtenus après la mise en place du plan de contrôle ne correspondent pas aux résultats prédits, on ne peut pas obligatoirement dire que

la qualité de l'échantillon n'était pas bonne. Le modèle statistique peut aussi être en cause ou bien de nouveaux facteurs ont pu apparaître. Il semble aussi que les modèles prédictifs, satisfaisants au niveau de la population, soient non satisfaisants pour les prédictions individuelles (Hadorn et coll. [35]). De nouvelles approches théoriques pourraient résoudre ce problème (Koopman et coll. [42]).

#### D - POST-MARKETING SURVEILLANCE - ESSAIS ECONOMIQUES

Les essais cliniques sont réalisés sur des échantillons souvent très différents des populations sur lesquelles les traitements sont ensuite utilisés. La pharmaco-épidémiologie propose d'effectuer, après les essais cliniques, des essais 'économiques' où les protocoles sont beaucoup moins stricts (eg. : pas d'étude à l'aveugle, le vétérinaire peut modifier le traitement quand il le souhaite, le diagnostic n'est pas confirmé, d'autres traitements peuvent être donnés en même temps que le produit à tester). Ce type d'essai, qui se déroule avant l'autorisation de mise sur le marché, pose cependant des problèmes éthiques et légaux. La surveillance post-marketing pourrait valider l'efficacité d'un traitement dans les conditions de terrain. Elle est cependant actuellement limitée à confirmer la sécurité du produit en enregistrant les réactions adverses.

"L'outcomes research" a été développée aux Etats-Unis pour évaluer l'efficacité des médicaments dans leurs conditions d'utilisation. La valeur de cette recherche est assez critiquée (Anderson, [2]).

#### E - EVALUATION DE LA PUISSANCE DES TESTS STATISTIQUES EFFECTUES

Le nombre d'unités statistiques échantillonnées doit être raisonné *a priori* en fonction des hypothèses testées, des puissances exigées des tests et des coûts du recueil des observations. La détermination du nombre d'unités statistiques est approchée car basée sur des estimations parfois peu précises des paramètres mesurés (prévalences et incidences généralement). Les tailles effectives des échantillons peuvent par ailleurs être inférieures aux nombres fixés au départ en raison, notamment, de refus de participer de la part de certains éleveurs ou d'abandons en cours d'étude.

Une fois l'enquête réalisée, il importe donc de vérifier si la taille des échantillons est suffisante (en terme essentiellement de puissance) pour détecter les effets des différents facteurs d'exposition étudiés (Rumeau-Rouquette et coll., [60]).

Une puissance insuffisante peut invalider complètement une étude.

Il est intéressant ici de remarquer que les enquêtes rétrospectives "cas/témoins" sont celles qui requièrent les tailles d'échantillons les plus petites à puissances égales (Fleiss, [30]).

En pratique, l'échantillonnage est soumis à de nombreuses contraintes qui ne sont pas toutes d'ordre méthodologique : les contraintes budgétaires ou les contraintes de temps sont souvent prépondérantes. Par ailleurs, s'il est aisé de définir par exemple la taille d'un échantillon lorsqu'on veut comparer deux échantillons pour un seul caractère, l'opération devient plus difficile, voire impossible si la comparaison porte sur plusieurs caractères. L'étude épidémiologique à visée exploratoire/explicative des maladies d'étiologie multifactorielle est particulièrement tributaire de la complexité des contraintes, et les enquêtes dites "d'écopathologie" en sont une bonne illustration.

Ces enquêtes vont maintenant être présentées à travers l'expérience acquise par le Laboratoire d'écopathologie de l'INRA-Theix et la Station de pathologie porcine du CNEVA - PLOUFRAGAN.

## SECONDE PARTIE : Pratique de l'échantillonnage

### I - PAR LE LABORATOIRE D'ECOPATHOLOGIE DE L'INRA-THEIX

Les enquêtes épidémiologiques analytiques réalisées par l'INRA\* se rattachent aux études dite "d'écopathologie" développées en France depuis le milieu des années 70 par Tillon et Madec dans la filière porcine, par Brochart et Barnouin dans la filière bovine, puis ultérieurement par l'équipe du Centre d'écopathologie animale (G. Rosner).

L'objectif principal de ces enquêtes est de mettre en évidence les facteurs de risque associés aux conditions d'environnement des troupeaux (bâtiment, climat, microbisme) et au fonctionnement technique de l'exploitation (alimentation, pratiques d'élevage). Au laboratoire d'écopathologie, les études concernent la vache laitière vivant dans des systèmes de production viables et les maladies d'élevage d'étiologie multifactorielle qui l'affectent pendant la durée moyenne d'une carrière de productrice et de reproductrice. Cela implique quelques contraintes majeures en matière d'échantillonnage :

- nécessité de prendre en compte deux échelles de perception : la vache et le troupeau ;
- nécessité de prendre en compte une durée d'étude optimale de 3-4 années et donc d'œuvrer avec des éleveurs fidèles et fiables ;
- nécessité d'intégrer la notion de système de production dans l'échantillon retenu.

Le laboratoire d'écopathologie a réalisé plusieurs enquêtes analytiques (EEPC, EEPB, SCEES, LASCAR) dans lesquelles les contraintes exposés ci-dessus ont guidé des stratégies d'échantillonnage à la fois différentes et complémentaires (Barnouin, [6]), (Barnouin, [7]), (Coulon et coll., [14]), (Faye et coll., [27]). L'unité statistique choisie a été le troupeau de vaches laitières productrices.

En fonction des objectifs choisis, ce troupeau comprenait la totalité des vaches productrices le composant, en incluant les génisses en fin de gestation (pour la recherche des facteurs de prédiction de la pathologie du peri-vêlage) et les veaux au cours du premier mois de leur existence (pathologies conséquentes aux troubles du vêlage affectant la mère). Le choix des animaux a donc été guidé par le choix des élevages. A ce titre, considérant que le système de production constitue une contrainte majeure de la santé du troupeau (il existe de fortes corrélations entre les différents types d'élevages et le profil de santé observable dans ces types) (Faye, [28]), l'échantillonnage prend en considération ces systèmes de façon prioritaire, d'autant plus qu'en termes opérationnels, le système de production détermine en grande partie la nature des mesures "anti-facteurs de risque".

On peut donc soit retenir des élevages en tenant compte de la diversité des systèmes de production rencontrés dans un espace géographique donné (cas de l'enquête écopathologique continue-EEPC : petits élevages intensifs du Finistère, grands élevages intensifs de Mayenne, petits élevages traditionnels d'Ardèche, élevages de montage de Savoie, grands élevages semi-intensifs de l'Est, ...), soit cibler l'échantillon sur un système de production donné (exemple de l'enquête écopathologique Bretagne-EEPB : élevages laitiers intensifs de Bretagne/race Holstein/production laitière moyenne économique > 6 500 kg en 1986 ...).

#### Choix de la taille de l'échantillon

La taille de l'échantillon est raisonnée en fonction de trois critères pratiques :

\* INRA : Institut National de la Recherche Agronomique



- La rareté des non-réponses et des abandons du fait du choix des éleveurs (cf. supra). On peut donc retenir dès le départ un échantillon relativement restreint,
- Le budget-temps qui permet de constater dans un échantillonnage par grappes (Frontier [31]) qu'il coûte plus cher d'augmenter le nombre d'élevages suivis que le nombre d'animaux (la prise en charge de 5-6 élevages complémentaires nécessite un enquêteur supplémentaire, alors qu'une augmentation de 15 % du nombre d'animaux ne modifie que très faiblement la charge globale de travail pour un nombre constant d'enquêteurs et n'affecte en termes de coût qu'un supplément associé aux éventuelles analyses de laboratoire). On oriente donc préférentiellement le choix vers moins d'élevages, mais avec de plus grands effectifs sachant que les estimations ainsi obtenues sont moins précises, à nombre total d'animaux fixé, que dans le cas d'un échantillonnage maximisant le nombre d'élevages (grappes). Le plan de sondage assure, ici un compromis entre précision d'échantillonnage et coût de recueil des observations,
- La prise en compte de la durée qui permet, sur certains critères de considérer les "élevages années" ou les « individus-lactations » comme les unités statistiques de base, ce qui permet de diminuer la taille de l'échantillon de départ.

Cette approche ne tient toutefois pas compte des corrélations temporelles (entre observations annuelles sur un même élevage) ni des corrélations spatiales (corrélations des observations entre animaux d'un même élevage).

Par ailleurs, les pathologies d'élevage auxquelles se consacrent généralement les études d'écopathologie sont présentes pratiquement dans tous les élevages et se posent rarement en termes de présence/absence.

Lorsque la sélection tient compte de la diversité des systèmes de production, la taille de l'échantillon doit être plus importante. En règle générale, ces diverses considérations conduisent au suivi d'un nombre d'élevages variant de 50 à plus d'une centaine selon la diversité des systèmes pris en compte et la durée du suivi. Dans le même temps, le nombre d'animaux suivis annuellement est de l'ordre de 2 à 4 000. A titre d'exemple, l'enquête écopathologique Bretagne a concerné 47 élevages pendant 4 ans (soit 188

élevages-années), 4 129 vaches laitières représentant 8 945 lactations complètes.

Compte tenu de la lourdeur du protocole, le choix des éleveurs est primordial pour éviter le biais qu'introduiraient les défections inévitables dans un échantillon aléatoire. La fiabilité des données est privilégiée et de ce fait la motivation de l'éleveur, son soin apporté à la notation des événements sont des clefs incontournables de sélection (Barnouin et coll., [4]). Par ce fait, les enquêtes d'écopathologie ont tendance à sélectionner des éleveurs plutôt tournés vers le technique, motivés par le sanitaire et par nature, consciencieux.

Cependant, pour assurer une variabilité suffisante du facteur à expliquer (la pathologie), dans un système de production homogène (EEPB), la consigne imposée aux partenaires chargés du choix des éleveurs (techniciens des organisations professionnelles, vétérinaires praticiens) est de présenter un couple d'exploitations, l'une avec peu de problèmes sanitaires en début d'étude, l'autre avec beaucoup de troubles d'étiologie mal élucidée.

Cette sélection des éleveurs dite à "choix raisonné" (ce que Thrusfield [68] appelle la sélection "intentionnelle" - *purposive selection*) nécessite, en toute rigueur, une validation des résultats.

Trois procédures ont été mises en place à l'INRA pour valider les résultats des enquêtes :

- Procédure de vérifications expérimentales quand cela est possible afin de tester les hypothèses étiologiques sous-jacentes aux facteurs de risque observés (étude du mécanisme du risque) : exemple du rôle de certains éléments de la ration alimentaire caractérisés par leur richesse en précurseurs des eicosanoïdes sur la pathologie du vêlage (rétention placentaire, fièvre vitulaire) (Barnouin et coll., [5]) ;
- Procédure de validation statistique sur un échantillon aléatoire proposé par le SCEES (Service Central des Etudes Economiques et Statistiques) : il s'agit d'un échantillon représentatif stratifié sur la base des finalités économiques dominantes des élevages (viande ou lait) et des effectifs du troupeau (3 strates) ; l'enquête de validation est une enquête prospective simplifiée : un simple questionnaire sur une feuille recto-verso rajoutée à l'enquête SCEES proprement dite qui permet de confirmer (ou d'infirmer à l'échelle de la population), les taux d'incidence des

pathologies observées dans l'enquête écopathologique et les facteurs de risque les plus importants identifiés à l'échelle du troupeau (Barnouin, [7]) ;

- Procédure de modélisation à partir des données rétrospectives et exhaustives des installations expérimentales de l'INRA (projet LASCAR). Ces données doivent permettre d'élaborer plusieurs modèles (fonctionnement des exploitations, gestion des carrières des

vaches laitières, prévision des maladies, effet des pathologies sur la production ...) qui seront validés sur les données collectées en situation réelle dans les enquêtes d'écopathologie (Lescourret et coll., [45]).

L'échantillonnage relève donc d'un ensemble de choix tendant à prendre en compte la complexité du fonctionnement et des relations individu-troupeaux qui caractérise les systèmes bovins laitiers.

## II - PAR L'UNITE DE RECHERCHES : « STATION DE PATHOLOGIE PORCINE » DU CNEVA - PLOUFRAGAN

### EXPOSE DU CONTEXTE DES ENQUETES ANALYTIQUES ET DE LA FAÇON DONT ON RAISONNE LA REPRESENTATIVITE

#### A - CARACTERISTIQUES DES TROUBLES PATHOLOGIQUES ETUDIES

Depuis une quinzaine d'années, des travaux d'épidémiologie sont conduits à la Station de Pathologie Porcine de Ploufragan. Ces travaux ont vu le jour à la suite d'un faisceau de circonstances parmi lesquelles l'incapacité de l'approche expérimentale exclusive à répondre à l'attente des partenaires de la filière notamment en matière de lutte préventive contre les maladies à fort impact économique. Les maladies à forte contagiosité (Grippe, SDRP, Aujeszky) ont essentiellement donné lieu à des travaux d'épidémiologie descriptive. En revanche, les maladies complexes, à caractère enzootique, ont été abordées selon une approche analytique où se fait la jonction avec l'écopathologie (Madec et coll. [48]). Il s'agit de rechercher les combinaisons de conditions associées aux différents niveaux de la pathologie étudiée. Par nature, en élevage considéré comme système ou ensemble organisé, les interactions entre facteurs sont en effet légion et on raisonne par conséquent préférentiellement en termes de profils de situation. Les troubles pathologiques

considérés présentent en commun quelques caractères :

- Le rôle de contaminants "microbiens" existe le plus souvent, mais ce rôle est loin d'être exclusif,
- Les caractéristiques physiques, physiologiques, immunologiques des animaux ainsi que les caractéristiques du milieu (alimentation, logement, ...) jouent un rôle majeur,
- Les troubles sont d'expression clinique progressive au point qu'on peut envisager une quantification et différents niveaux de sévérité. Le plus souvent, la mortalité est réduite mais la morbidité est élevée. Ce sont des "maladies du quotidien" dans les élevages,
- Enfin, la reproduction expérimentale de ces pathologies est difficile. Elle nécessite de recourir à des méthodes particulières et à des modèles expérimentaux, conditions fort éloignées des conditions d'éclosion spontanée de la pathologie dans les élevages.

Toutes ces caractéristiques ne sont pas dénuées de conséquences quant à la démarche à adopter pour atteindre le but fixé. Elles nous conduisent en effet à mener des investigations dans les élevages eux-mêmes. Pour l'entité pathologique, objet de l'étude, ces investigations consistent à réaliser une description ciblée des animaux, du milieu ainsi que des faits pathologiques à l'aide d'une collection de paramètres descripteurs élaborés préalablement à l'enquête (Madec [47]).

La réalisation d'une description objective entraîne le plus souvent des manipulations standardisées sur les animaux (prises de sang, pesées, examens cliniques, ...) et parfois l'achat d'animaux pour autopsie et recherches de laboratoire. Enfin, les enquêtes prospectives sont permises en raison de la brièveté des cycles biologiques de l'espèce porcine. Les élevages mis en observation sont ainsi l'objet de visites régulières au long de l'enquête dont la durée est en général de 1 à 6 mois.

### B - CARACTERISTIQUES DE L'ORGANISATION DE LA PRODUCTION

Les élevages porcins sont en grande majorité affiliés à des structures à vocation technique, commerciale ou industrielle. Ces dernières sont pour la plupart dotées d'un encadrement technique et vétérinaire qui assure le suivi sanitaire des élevages. La visée pragmatique des travaux d'épidémiologie est guidée par la demande de repères fiables permettant d'asseoir des stratégies de lutte préventive. Les études sont parfois cofinancées de manière directe ou indirecte par les groupes économiques : firmes, services, coopératives. Leurs représentants sont associés au processus d'élaboration du protocole et de collecte des données.

### C - CONSEQUENCES PRATIQUES POUR LE CHOIX DES ELEVAGES

Il découle des lignes précédentes la nécessité du volontariat de l'éleveur. Ce dernier, présent au quotidien dans l'élevage est en outre un observateur de choix. Le travail ne peut raisonnablement se dérouler que chez des éleveurs qui acceptent de participer activement à l'étude et qui donc acceptent les contraintes du protocole. Il faut également l'aval de la structure économique à laquelle se rattache l'élevage, d'autant que celle-ci est impliquée dans le processus. Enfin, une compatibilité enquêteur-enquêté est nécessaire en raison de la nature et de la durée des travaux. De l'harmonie et l'efficacité du binôme enquêteur-enquêté dépend la qualité des données collectées. L'importance accordée à cette harmonie ainsi qu'à l'ensemble des travaux préparatoires aux enquêtes est également exprimée dans des rapports concernant d'autres espèces animales étudiées en France (Philippot, [54]), (Calavas, [12]).

### D - DEFINITION DES CONTOURS DE LA POPULATION-CIBLE ET AUTRES CONDITIONS DU RECRUTEMENT

Les conditions techniques générales de la production porcine ont convergé au cours des dernières décennies vers une harmonisation aboutissant à la mise en place de systèmes structurés de façon relativement homogène. De même, on assiste à une rationalisation des techniques de conduite de ces systèmes. Ainsi est apparu le système confiné intensif qui fournit aujourd'hui la grande majorité des porcs français. Le plus souvent, les contours de la population-cible se dessinent naturellement dès la définition du problème soumis à investigation. Une étude sur une pathologie des reproducteurs élevés en plein air conduira d'emblée à travailler en système plein air à l'exclusion du système confiné. De même, les élevages peuvent être rangés selon la fonction qu'ils assurent dans la chaîne de production de viande et celle-ci intervient dans le choix. Ainsi, une étude sur truies et/ou sur porcelets exclura les élevages spécialisés dans la phase d'engraissement (élevages "engraisseurs").

On peut donc d'ores et déjà conclure que le recrutement des élevages se fait dans un système général de production homogène. Par ailleurs, il y a lieu de travailler dans une zone géographique aussi homogène que possible. Ce point est d'importance particulière pour certaines études notamment en système ouvert (élevages en plein air). A l'intérieur du cadre ainsi tracé une grande variabilité est, en revanche, recherchée dans l'intensité du phénomène pathologique étudié. Celle-ci ne peut être qu'approximée car évaluée *a priori* par les responsables technico-sanitaires qui assurent le suivi habituel des élevages.

Par ailleurs, le traitement des données est orienté vers la comparaison de groupes d'unités statistiques présentant le problème à différents niveaux d'une échelle (depuis l'absence de problème jusqu'au problème très sévère). Une dispersion insuffisante sur l'échelle nuit singulièrement à la recherche ultérieure des circonstances associées aux différents niveaux (facteurs de risque). Aussi est-il demandé d'emblée aux enquêteurs de solliciter des éleveurs couvrant l'ensemble du spectre, en particulier les extrêmes.

## E - LA NOTION DE REPRESENTATIVITE

La représentativité d'un échantillon donne des garanties de validité en matière d'inférence c'est-à-dire d'extrapolation des résultats à l'ensemble de la population dont il est extrait. A partir des propos tenus au paragraphe précédent, on constate clairement que l'échantillon doit être aussi proche que possible de la population cible pour ce qui concerne les conditions générales d'exploitation des animaux. En revanche, cette proximité n'est pas recherchée pour ce qui touche à l'intensité de la pathologie étudiée (variable à expliquer). Faut-il rappeler que le but de ces enquêtes n'est pas d'estimer une prévalence dans la population-cible à partir de l'échantillon. Le tirage au sort est cependant à recommander pour des enquêtes peu contraignantes. Nous venons d'ailleurs d'y recourir pour une enquête transversale conduite conjointement avec l'INRA et l'ITP\* (2 visites à l'élevage sur 2 jours consécutifs, les maîtres d'oeuvre étant également enquêteurs exclusifs). La population d'où provient l'échantillon est celle des élevages possédant des truies et pratiquant la gestion technique des troupeaux (GTTT, base nationale de plus de 4 000 troupeaux). En revanche, dès que les contraintes en tout genre s'accumulent, le tirage au sort n'est en général ni accessible ni même souhaitable pour le recrutement, car au-delà des risques de refus et de pertes de temps, la qualité des données peut se trouver affectée (exhaustivité, véracité, ...). Ainsi, un choix raisonné est-il pratiqué le plus souvent. Pour ces études, la conformité de l'échantillon à la population-cible en regard des conditions générales d'exploitation ne peut être qu'approchée et obtenue en fin d'étude.

□ La première difficulté est celle du choix des paramètres descripteurs à prendre en compte pour juger de la "ressemblance" *a posteriori* entre les élevages de l'échantillon et ceux de la population-

cible. A cet égard il est raisonnable de considérer :

- des critères de structure d'élevage. Ils décrivent la configuration de l'outil de travail,
  - des critères de conduite. Ils fournissent les grandes lignes des techniques employées pour utiliser l'outil,
  - ces critères peuvent être complétés par des éléments de résultats. Toutefois, certains de ces derniers étant naturellement liés à la pathologie étudiée, la vigilance est de mise dans l'interprétation.
- La seconde difficulté est de disposer de registres à la fois ciblés sur la population et suffisamment exhaustifs et actuels pour réaliser les comparaisons

Enfin, aucun redressement de l'échantillon n'est réalisé *a posteriori*. Au moment de la publication des résultats, l'échantillon, support du travail, est soigneusement décrit. Une déviation excessive conduit à des consignes de prudence particulière lors de l'utilisation des résultats dans des élevages qui fonctionnent selon un mode différent de celui du groupe considéré.

Quoiqu'il en soit, au terme de l'enquête, une étape de validation des facteurs de risque est réalisée. Elle consiste dans des élevages de la population-cible affectés par la pathologie étudiée :

- à évaluer la situation à l'égard des facteurs de risque issus de l'enquête,
- à mettre en place des mesures de correction des facteurs de risque défaillants,
- à mesurer l'impact de ces mesures nouvelles sur le phénomène étudié.

\* ITP : Institut Technique du Porc

**LE CAS D'UNE ENQUETE :  
LA PATHOLOGIE RESPIRATOIRE  
ENZOOTIQUE  
CHEZ LE PORC EN PHASE  
D'ENGRASSEMENT  
(ANNEE 1980, SPP 1981)**

### A - LES TRAVAUX PRELIMINAIRES

Les travaux préliminaires aux enquêtes ont pour but la mise au point de paramètres descripteurs et plus généralement l'acquisition d'expertise pour le maître d'œuvre, le tout conduisant à un protocole adapté. Ces travaux ont permis de suspecter l'influence d'événements antérieurs à la phase d'engrassment dans l'expression de la pathologie étudiée. Ainsi, avons-nous décidé de limiter l'étude aux élevages pratiquant à la fois le naissage et l'engrassment (élevages naisseurs-engrassment) à l'exclusion des "engrassment". En 1980, environ 60 % des porcs à l'engrais en Bretagne étaient élevés dans de telles exploitations. La population cible est définie comme l'ensemble des élevages naisseurs-engrassment organisés de la région Bretagne pratiquant en système intensif confiné. Par élevage organisé on entend des unités qui mettent en œuvre une méthode de travail : conduite en lots (bandes), identification des reproducteurs, tenue de documents d'élevage, ...

### B - LE RECRUTEMENT

Une latitude est laissée aux enquêteurs pour le choix des modèles de bâtiments (élevages ouverts cependant exclus). En revanche, au sein des groupes économiques il est demandé de choisir des élevages ayant *a priori* un statut sanitaire varié en regard de la pathologie étudiée en insistant sur les extrêmes.

### C - RESULTATS

Un échantillon de 103 élevages est obtenu. Au total, 7500 porcs sont suivis depuis la naissance jusqu'à l'abattage (durée de l'étude : 160 à 220 jours selon l'élevage). Au terme de l'étude, l'échantillon est comparé aux données de

registres comportant parfois plusieurs centaines de troupeaux (Tableau II).

- ♦ Deux éléments de structure sont retenus : la taille des élevages et le mode de renouvellement du troupeau. Pour le critère taille, le groupe de référence choisi est l'ensemble des élevages pratiquant la gestion technique, c'est-à-dire la démarche d'enregistrement des performances de reproduction (Dagorn [16]). Selon notre expertise, la fraction des "naisseurs-engrassment" du groupe est très proche de la population cible. La taille moyenne des élevages enquêtés est de 75 truies contre 63 pour le groupe de référence. On en déduit que l'enquête s'est déroulée dans des élevages de taille supérieure.

Le second critère est l'origine des reproducteurs de renouvellement. Le fichier précédent ne fournit pas cette information. Nous nous sommes donc basés sur une étude descriptive des structures de la production réalisée par l'Institut Technique du Porc à la même époque (ITP [38]). Ces élevages (N = 430) en Bretagne sont un échantillon aléatoire des élevages pratiquant des enregistrements-types de performances. On observe ici encore une sensible déviation. Le groupe de travail pratique moins l'autorenouvellement et s'approvisionne par conséquent davantage à l'extérieur.

- ♦ Trois éléments généraux de la conduite d'élevage sont retenus. Deux concernent les conditions de logement des porcs à l'engrais, catégorie d'animaux sur laquelle porte l'enquête. Le groupe de travail diffère du groupe de référence en ce qui concerne le mode de ventilation de la porcherie d'engrassment. Pour le troisième critère, l'âge au sevrage, les valeurs sont identiques.

- ♦ Deux éléments de performances de troupeaux sont mentionnés au Tableau 2. Le nombre de porcelets sevrés par truie et par an ou productivité numérique est un critère synthétique très largement utilisé. Il est voisin dans les deux groupes. Enfin, pour l'indice de classement des carcasses à l'abattoir le groupe des élevages enquêtés est comparé à un groupe d'élevages naisseurs-engrassment pratiquant la gestion technicoéconomique (Dagorn [15]) seule source disponible. Le niveau des performances (moyenne et écart-type) est très proche dans les deux groupes.

Tableau II : Comparaison rétrospective de l'échantillon des élevages enquêtés par rapport à différents groupes de référence  
(1) Dagorn et coll. 1981 ; (2) ITP 1982 ; (3) Dagorn 1981 ; (NS : non significatif)

Paramètres choisis	Echantillon (N = 103)	Groupe de référence		Signification
- Taille de l'élevage nb de truies ; moy. (écart-type)	75 (50)	63 (42)	N = 2120 <sup>(1)</sup>	p<0.001
- Origine des reproducteurs en %				
• achat extérieur à 100 %	45	34		
• autorenouvellement	10	17		
• mixte pour truies	45	49	N = 430 <sup>(2)</sup>	p<0.05
- Type de sol en engraissement (%)				
• caillebotis intégral	27	26		
• caillebotis partiel	68	65		
• autres	5	9	N = 430	NS
- % places d'engraissement en ventilation dynamique	48	33	N = 430	p<0.001
- Age au sevrage des porcelets (j.)	30	30	N = 2120	NS
- Nb de porcelets sevrés/truie/an moy. ; (écart-type)	20.0 (2.2)	20.1 (2.4)	N = 2120	NS
- Indice classement des carcasses (selon grille en vigueur) moy. ; (écart-type)	2.3 (0.3)	2.4 (0.3)	N = 854 <sup>(3)</sup>	NS

La lecture du tableau II montre globalement que les élevages enquêtés sont sensiblement de plus grande taille et mettent en œuvre des techniques plus sophistiquées que les élevages pris pour comparaison. Le recul du temps nous fait dire aujourd'hui que les élevages suivis étaient plutôt aux avant-postes de l'intensification et de la modernisation. En effet, en 1994 la taille moyenne des élevages est nettement plus élevée qu'en 1980, l'achat à l'extérieur des

reproducteurs de renouvellement est la règle et des ventilateurs sont en place dans l'immense majorité des porcheries d'engraissement. Aussi, à défaut d'une "représentativité", d'accès très difficile et probablement illusoire, nous disions à l'époque que "l'échantillon constituait une bonne base de travail pour les études épidémiologiques sur des pathologies complexes accompagnant l'intensification" (SPP [67]).

## CONCLUSION

En médecine vétérinaire, beaucoup d'études épidémiologiques analytiques ont été faites, ces quinze dernières années, pour mieux comprendre les pathologies multi-factorielles. Si la méthodologie statistique s'est développée en parallèle pour mieux analyser les résultats, moins d'efforts ont été consacrés à la méthodologie d'échantillonnage et à la vérification de la qualité de l'échantillon.

Soit l'échantillon est tiré de façon aléatoire et l'on rencontre alors des problèmes de refus de participation ou alors les variables enregistrées

sont de faible qualité par rapport au problème biologique étudié ; soit l'échantillon est sélectionné par choix raisonné et se pose alors le problème de l'extrapolation des résultats.

Pourtant ces problèmes sont peu ou pas évoqués dans les discussions des articles publiant ces études épidémiologiques alors qu'ils contribuent aussi de façon essentielle à la validité des conclusions tirées. Quelques méthodes de validation de la qualité de l'échantillon ont été développées et certaines déjà utilisées en médecine vétérinaire.

## BIBLIOGRAPHIE

- [1] Akhtar S., Gardner I.A., Hird D.W. and Holmes J.C. - Computer simulation to compare three sampling plans for health and production surveillance in California dairy herds. *Prev. Vet. Med.* 1988, **6**, 171-181.
- [2] Anderson C. - Measuring what works in health care. *Science* 1994, **263**, 1080-1082.
- [3] Andersson L. and Emanuelson U. - An epidemiological study of hyperketonaemia in swedish dairy cows ; determinants and the relation to fertility. *Prev. Vet. Med.* 1985, **3**, 449-462.
- [4] Barnouin J. et Brochart M. - Enquête écopathologique continue : les objectifs et leur réalisation, le choix et la typologie des élevages. *Ann. Rech. Vet.* 1986, **17** (3), 201-207.
- [5] Barnouin J. and Chassagne M. - An aetiological hypothesis for the nutrition-induced association between retained placenta and milk fever in the dairy cow. *Ann. Rech. Vet.* 1991, **22**, 331-343.
- [6] Barnouin J. - Enquête écopathologique continue en élevages-observatoires chez les ruminants : objectifs et stratégie. *Ann. Rech. Vet.* 1980, **11**, 341-350.
- [7] Barnouin J. - Pathologie des vaches laitières en Bretagne. Relations avec l'alimentation et le logement. *Cah. Stat. Agric.* 1988, **2** (6), 35-48.
- [8] Bauer P. - Two stage sampling for simultaneously testing main and side effects in clinical trials. *Biom. J.* 1986, **28**, 871-879.
- [9] Bendixen P.H., Vilson B., Ekesbo I. and Astrand D.B. - Disease frequencies of tied zero-grazing dairy cows and of dairy cows on pasture during summer and tied during winter. *Prev. Vet. Med.* 1986, **4**, 291-306.
- [10] Bigras-Poulin M., Meek A.H. and Martin S.W. - Interrelationships of health problems and age on milk production in selected Ontario Holstein cows.
- [11] Blackwelder W.C. - Sample size and power for prospective analysis of relative risk. *Stat. Med.* 1993, **12**, 691-698.

- [12] Calavas D. - Mammites des brebis allaitantes. Compte rendu d'enquête n° 11. 1992 Centre d'écopathologie animale. Villeurbanne.
- [13] Chen C.H. and George S.L. - The bootstrap and identification of prognostic factors via Cox's proportional hazards regression model. *Stat. Med.* 1985, 4, 39-46.
- [14] Coulon J.B., Lescourret F., Faye B., Landais E., Troccon J.L. et Perochon L. - Description de la base de données "LASCAR", un outil pour l'étude des carrières des vaches laitières. *INRA Prod. Anim.* 1993, 6 (2), 151-160.
- [15] Dagorn J. - La gestion technique des troupeaux de truies. Résultats départementaux pour l'année 1980. *Techni porc* 1981, 4 (4) 31-37.
- [16] Dagorn J., De Jubecourt M., Le Tiran M.M., Owen J. et Runavot J.P. - Les performances de l'année 1980. *Techni porc* 1981, 4 (3) 25-32.
- [17] Day S.J. and Graham D.F. - Sample size estimation for comparing two or more treatment groups in clinical trials. *Stat. Med.* 1991, 10, 33-43.
- [18] Dohoo I.R., Martin S. W., Meek A.H. and Sandals W.C.D. - Disease, production and culling in Hostein-Friesian cows. I - The data. *Prev. Vet. Med.* 1983, 1, 321-334.
- [19] Dohoo I.R., Martin S. W., Mc Millan I. and Kennedy B.W. - Disease, production and culling in Hostein-Friesian cows. II - Age, season and sire effects. *Prev. Vet. Med.* 1984, 2, 655-670.
- [20] Dohoo I.R. and Martin S. W. - Disease, production and culling in Hostein-Friesian cows. III - Disease and production as determinants of disease. *Prev. Vet. Med.* 1984, 2, 671-690.
- [21] Dohoo I.R. and Martin S. W. - Disease, production and culling in Hostein-Friesian cows. IV - Effects of disease on production. *Prev. Vet. Med.* 1984, 2, 755-770.
- [22] Dohoo I.R. and Martin S. W. - Disease, production and culling in Hostein-Friesian cows. V - Survivorship. *Prev. Vet. Med.* 1984, 2, 771-784.
- [23] Donner A. - Approches to sample size estimations in the design of clinical trials. A review. *Stat. Med.* 1984, 3, 199-214.
- [24] Donner A., Brown K.S. and Brasher P. - A methodological review of non-therapeutic intervention trials employing cluster randomization. *Int. J. Epidemiol.* 1990, 19, 795-800.
- [25] Donner A. - Sample size requirements for stratified cluster randomization designs. *Stat. Med.* 1992, 11, 743-750.
- [26] Efron B. - The Jackknife, the Bootstrap and other resampling plans. Society for Industrial and Applied Mathematics, 1982.
- [27] Faye B., Barnouin J. et Lescourret F. - Objectifs principaux et stratégie de l'enquête écopathologique Bretagne sur la vache laitière. *Epidémiol. Santé Anim.* 1989, 15, 23-31.
- [28] Faye B. - Interrelationship between health status and farm management system in french dairy herds. *Prev. Vet. Med.* 1991, 12, 133-152.
- [29] Flanders W.D. and Greenland S. - Analytic methods for two-stage case-control studies and other stratified designs. *Stat. Med.* 1991, 10, 739-747.
- [30] Fleiss J.L. - Statistical methods for rates and proportions. 1973, Toronto, Canada: John Wiley & Sons.
- [31] Frontier S. - Stratégies d'échantillonnage en écologie. Ed. Masson, 1983, Paris, 494 p.
- [32] Gardner I.A., Hird D.W., Utterback W.W., Danaye-Elmi C., Heron B.R., Christiansen K.H. and Sischo W. M. - Mortality, morbidity, case-fatality and culling rates for California dairy cattle as evaluated by the National Animal Health Monitoring System, 1986-1987. *Prev. Vet. Med.* 1990, 8, 157-170.
- [33] Graubard B.I., Fears T.R. and Gail M.H. - Effects of cluster sampling on epidemiologic analysis in population-based case-control studies. *Biometrics* 1989, 45, 1053-1071.
- [34] Greenland S. - Power, sample size and smallest detectable effect determination for multivariate studies. *Stat. Med.* 1985, 4, 117-127.



- [35] Hadorn D.C., Draper D., Rogers W.H., Keeler E.B. and Brook R.H. - Cross-validation performance of mortality prediction models. *Stat. Med.* 1992, 11, 475-489.
- [36] Hird D.W., Weigler B.J., Salman M.D., Danaye-Elmi C., Palmer C.W., Holmes J.C., Utterback W.W. and Sischo W.M. - Expenditures for veterinary services and other costs of disease and disease prevention in 57 California beef herds in the National Animal Health Monitoring System, 1988-1989. *JAVMA* 1991, 4, 554-558.
- [37] Hurd H.S. and Kaneene J.B. - The National Animal Health Monitoring System in Michigan. II - Methodological issues in the estimation of frequencies of disease in a prospective study of multiple dynamic populations. *Prev. Vet. Med.* 1990, 8, 115-125.
- [38] ITP - Enquête Gestion Technique des troupeaux de truies. Descriptifs régionaux. Document ITP 1982, 57 pages + tableaux.
- [39] Kaneene J.B. and Hurd H.S. - The National Animal Health Monitoring System in Michigan. I - Design, data and frequencies of selected dairy cattle diseases. *Prev. Vet. Med.* 1990, 8, 103-114.
- [40] Kaneene J.B. and Hurd H.S. - The National Animal Health Monitoring System in Michigan. III - Cost estimates of selected dairy cattle diseases. *Prev. Vet. Med.* 1990, 8, 127-140.
- [41] Koepsell T.D., Martin D.C., Diehr P.H., Psaty B.M., Wagner E.H., Perrin E.B. and Cheadle A. - Data analysis and sample size issues in evaluations of community-based health promotion and disease prevention programs : a mixed-model analysis of variance. *J. Clin. Epidemiol.* 1991, 44, 701-713.
- [42] Koopman J.S. and Weed D.L. - Epigenesis theory : a mathematical model relating causal concept of pathogenesis in individuals to disease patterns in populations. *Am. J. Epidemiol.* 1990, 132, 366-390.
- [43] Lakatos E. and Gordon Lan K.K. - A comparison of sample size methods for logrank statistic. *Stat. Med.* 1992, 11, 179-191.
- [44] Lemeshow S., Hosmer Jr D.W. and Klar J. - Sample size requirements for studies estimating odds ratios or relative risks. *Stat. Med.* 1988, 7, 759-764.
- [45] Lescourret F., Perochon L., Coulon J.B., Faye B. and Landais E. - Modelling an information System using the MESISE method for Agricultural Research : the example of a database for a study on performances in Dairy cows. *Agric. Syst.* 1992, 38, 149-173.
- [46] Lubin J.H. and Gail M.H. - On power and sample size for studying features of the relative odds of disease. *Am. J. Epidemiol.* 1990, 131, 552-566.
- [47] Madec F. - Travaux préalables à une enquête écopathologique : la recherche de paramètres descripteurs. *Ann. Med. Vet.* 1994, 138, 55-63.
- [48] Madec F. et Tillon J.P. - Ecopathologie et facteurs de risque en Médecine Vétérinaire. Analyse rétrospective de l'expérience acquise en élevage porcin intensif. *Rec. Med. Vet.* 1988, 164, 607-616.
- [49] Martin S.W., Meek A.H. and Willeberg P. - Veterinary epidemiology, principles and methods. 2d edition 1988, Iowa State University Press, Ames, Iowa.
- [50] Mc Dermott J.J., Alves D.M., Anderson N.G. and Martin S.W. - "Benchmark" - a large observational study of Ontario beef breeding herds : study design and collection of data. *Can. J. Vet. Res.* 1991, 32, 407-412.
- [51] Mc Dermott J.J., Alves D.M., Anderson N.G. and Martin S.W. - Measures of herd health and productivity in Ontario cow-calf herds. *Can. J. Vet. Res.* 1991, 32, 413-420.
- [52] Meek A.H., Martin S.W., Stone J.B., Mc Millan I., Britney J.B. and Grieve D.G. - The relationship among current management systems, production, disease and drug usage on Ontario dairy farms. *Can. J. Vet. Res.* 1986, 50, 7-14.
- [53] Meydrecht E.F. and Kupper L.L. - Cost considerations and sample size requirements in cohort and case-control studies. *Am. J. Epidemiol.* 1978, 107:201-5.

- [54] Philippot J.M. - Fourbure et fourchet des vaches laitières. Compte rendu d'enquête n° 8. 1992 Centre d'écopathologie animale. Villeurbanne.
- [55] Philippot J.M., Pluvinage P., Cimarosti I. and Luquet F. - On indicators of laminitis and heelhorn erosion in dairy cattle : a research based on the observation of digital lesions, in the course of an ecopathological survey. *Proceedings of the VII<sup>th</sup> international symposium on diseases of the ruminant digit, update in cattle lameness*. Edition Schering-Plough Liverpool July 1990, 184-198.
- [56] Pike M.C. and Casagrande J.T. - Re: Cost considerations and sample size requirements in cohort and case-control studies. *Am. J. Epidemiol.* 1979, 110:110-2.
- [57] Pluvinage P., Ducruet T., Josse J. et Monicat F. - Facteurs de risque des mammites des vaches laitières. *Rev. Med. Vet.* 1991, 167, 105-112.
- [58] Reinke W.A. - Applicability of industrial sampling techniques to epidemiologic investigations : examination of an underutilized resource. *Am. J. Epidemiol.* 1991, 134, 1222-1232.
- [59] Royston P. - Exact conditional and unconditional sample size for pair-matched studies with binary outcome : a practical guide. *Stat. Med.* 1993, 12, 699-712.
- [60] Rumeau-Rouquette C., Bréart G. et Padiou R. - Méthodes en épidémiologie. 1981, Flammarion, Médecine-Science.
- [61] Sauerbrei W. and Schumaber M. - A bootstrap resampling procedure for model building : application to the Cox regression model. *Stat. Med.* 1992, 11, 2093-2109.
- [62] Schoenfeld D.A. - Sample-size formula for the proportional-hazards regression model. *Biometrics* 1983, 39, 499-503.
- [63] Schwartzbaum J.A. and Hirschberg J.G. - Exploring the order of odds ratios using the Bootstrap. *Stat. Med.* 1991, 10, 1273-1281.
- [64] Shoukri M.M. and Martin S.W. - Estimating the number of clusters for the analysis of correlated binary response variables from unbalanced data. *Stat. Med.* 1992, 11, 751-760.
- [65] Sischo W.M., Hird D.W., Gardner I.A., Utterback W.W., Christiansen K.H., Carpenter T.E., Danaye-Elmi C. and Heron B.R. - Economics of disease occurrence and prevention on California dairy farms : a report and evaluation of data collected for the National Animal Health Monitoring System, 1986-1987. *Prev. Vet. Med.* 1990, 8, 141-156.
- [66] Smith J., Connett J. and Mc Hugh R. - Planning the size of a matched case-control study for estimation of the odds ratio. *Am. J. Epidemiol.* 1985, 122, 345-347.
- [67] SPP - Application d'une méthode d'étude sanitaire globale à la prévention des troubles respiratoires chez le porc à l'engrais. Document Station de Pathologie Porcine. 1981, 85 pages + annexes.
- [68] Thrusfield M. - Veterinary epidemiology. Ed. Butterworths, 1986, London, 280 p.
- [69] Weinberg C.R. and Sandler D.P. - Randomized recruitment in case-control studies. *Am. J. Epidemiol.* 1991, 134, 421-432.
- [70] Yateman N.A. and Skene A.M. - Sample sizes with arbitrary patient entry and loss to follow-up distributions. *Stat. Med.* 1992, 11, 1103-1113.