

LA GESTION DE L'INFORMATION DANS LES ENQUÊTES D'ÉPIDÉMIOLOGIE ANIMALE

P. Sulpice ^[1], Françoise Lescourret ^[2], D. Calavas ^[1],
Françoise Bugnard ^[1] et B. Faye ^[2],

Résumé

L'information exploitée dans les enquêtes d'épidémiologie animale est de nature complexe, souvent peu formalisée a priori et d'origine variée. Sa gestion, considérée comme l'élément fondamental d'un processus qui permet à une étude d'aboutir, doit être imaginée et réalisée simultanément à la conception de l'enquête. Elle se définit par l'organisation du recueil des données et par le contrôle de leur qualité (c'est-à-dire le contrôle de la qualité des opérations qui produisent les données), tout au long du déroulement de l'enquête. Ces deux opérations, l'organisation du recueil et le contrôle de la qualité qui correspondent à la gestion de l'information, sont indispensables à l'obtention de données fiables, étape incontournable pour mener à bien une enquête d'épidémiologie animale. Cet article présente des méthodes et des moyens nécessaires à réunir pour obtenir une gestion de l'information adaptée à de tels objectifs.

Summary

Data used in animal epidemiology surveys are complex, not always a priori well formalized, and they have various origins. Their management is one of the key point of a good standard study; it has to be organised and realised from the beginning of the study planning. The data management involves, throughout the study, the data collection planning and the checking of the data quality (i.e. to check the different procedures used to produce the data). These two aspects of the data management are both required to obtain reliable and accurate data, critical point to perform good quality epidemiological research. This paper describes some methods and means usefull to fulfill the goals of a good data management.



L'enquête peut être définie comme la recherche méthodique d'information reposant sur des mesures, des questions et des témoignages [Toma et al., 1991]. Cette information, recueillie en fonction des objectifs assignés à l'enquête (enquêtes descriptive, analytique, finalisée ou à visée de recherche exploratoire, ou d'évaluation)

est à la fois la matière première (l'information "collectée") et le produit fini (l'information "transformée") de l'enquête d'épidémiologie animale. Pour passer du premier au deuxième stade elle doit être collectée, traitée et conservée : elle est gérée.

[1] Centre d'Ecopathologie animale, 26 rue de la Baisse, 69100 Villeurbanne, France

[2] Laboratoire d'Ecopathologie, INRA Theix, 63122 St Genès Champanelle, France

L'acquisition et l'organisation des informations depuis leur lieu initial (leur gisement) jusqu'à celui de la base de données, qui permet de les organiser et de les structurer [Hall et al., 1980 ; Stone et Thrusfield, 1989] et les moyens humains et organisationnels qui contribuent à l'obtention de l'information (la gestion des relations humaines qui rend possible la collecte par des acteurs d'origine sociale et professionnelle hétérogène : éleveur, enquêteur, technicien... ; la gestion administrative : conventions d'enquête... ; la gestion organisationnelle : matériel, questionnaires, diversité des modes d'acheminement) relèvent de la gestion de l'information.

Celle-ci est une préoccupation constante dans les enquêtes d'épidémiologie animale [Barnouin, 1980 ; Josse, 1986]. La qualité de l'information "collectée" conditionne la qualité de l'information "transformée", la donnée ; elle conditionne la qualité globale de l'enquête. Comme dans tout processus de production, un contrôle de la qualité doit être mis en place tout au long de la chaîne de production et de gestion pour assurer la qualité du produit final [Faye et al., 1994].

L'épanouissement considérable des possibilités informatiques intervenu ces dernières années, tant au niveau du matériel qu'à celui des logiciels, permet maintenant de construire des procédures de gestion du recueil des données et de contrôle de la qualité.

I - L'INFORMATION DANS LES ENQUÊTES D'ÉPIDÉMIOLOGIE ANIMALE

L'information exploitée dans les enquêtes d'épidémiologie animale (E.E.A.) est difficile à gérer, en raison de la diversité de sa nature (mesures, appréciation) et de son origine (élevages, organismes professionnels et économiques, laboratoires...), de sa complexité et de son volume.

A - NATURE

L'information se présente sous différents formats : numériques (entiers et réels), alphanumériques (nom usuel des animaux, matricule d'identification) ainsi que des dates ou des images (imagerie médicale).

L'enquête amène souvent à décrire une information complexe et difficile à cerner (pathologie, hygiène, ambiance d'un bâtiment, état d'engraissement). Il est nécessaire d'avoir recours à des "indicateurs" (échelles de notation, appréciations semi-quantitatives) et des "définitions alternatives" [Curtis, 1989], par

exemple, pour les pathologies, des mesures physiologiques ou des mesures de performances à la place de signes cliniques. Cela se traduit par une standardisation au prix d'une réduction de l'information. L'information collectée est une image appauvrie mais commune de la réalité.

L'information produite par un événement ou une observation peut être de nature pérenne ou furtive comme lorsqu'elle est collectée à une date ou à un stade physiologique précis (cf. encart 1).

Son niveau de formalisation *a priori* est variable et en général faible [Calavas et al., 1994]. En effet, nombreuses sont les données qui sont recueillies spécifiquement pour les besoins de l'enquête, et pour lesquelles il n'existe aucun enregistrement en routine (par exemple, information qui n'est pas liée à une obligation légale ou qui n'est pas relevée dans le cadre de la gestion de l'élevage, description très détaillée d'un événement).

Encart 1**INFORMATIONS PERENNES ET INFORMATIONS FURTIVES**

La pérennité d'une information dépend de son degré de formalisation *a priori* qui est lui-même lié à son caractère utile ou obligatoire (information habituellement relevée pour la conduite de l'élevage, pour répondre à des obligations légales). Elle dépend également de la volatilité du support sur lequel elle est enregistrée (mémoire de l'éleveur, document papier, support informatique), de la possibilité de produire à nouveau l'information et de sa stabilité dans le temps (la surface d'une étable est une information pérenne). La date de vêlage est, par exemple, pérenne pour un éleveur adhérent à un système de contrôle de performances (Contrôle laitier). En revanche, un état d'engraissement en fin de gestation est l'archétype d'une information furtive.

L'importance d'une information dans une enquête d'épidémiologie animale est (malheureusement) sans lien avec son caractère pérenne/furtif, les informations nécessaires à la construction des variables à expliquer dans une étude étant fréquemment furtives (pathologiques, caractéristiques à un stade physiologique donné...).

En conséquence, il est nécessaire de choisir et de mettre en oeuvre la procédure de recueil adaptée au caractère pérenne/furtif de l'information que l'on souhaite recueillir.

En effet, les informations pérennes sont récupérables par des relances en cas de « données manquantes », ou d'erreur sur les données lors du premier relevé d'information. En revanche, pour les informations furtives (notamment les enregistrements à effectuer à une date ou un stade physiologique précis), des protocoles précis doivent être mis en place avec, par exemple, des outils alertant l'enquêteur sur les mesures à faire à telle date ou à tel stade physiologique afin de limiter une absence irrémédiable donnée.

B - ORIGINE

Dans les élevages, les informations sont en général collectées à la fois par l'éleveur et par l'enquêteur. En plus de ces deux intervenants, on peut avoir recours parfois à un enquêteur "secondaire" chargé de la collecte d'une information spécifique en rapport avec son activité professionnelle sur le site (inséminateur, technicien d'abattoir, vétérinaire sanitaire...) [Philipot, 1993].

D'autre part, les E.E.A. font fréquemment appel à d'autres sources d'information : organisations professionnelles agricoles (O.P.A.) comme les organismes de contrôle de performances et les centres d'insémination artificielle, Direction des Services vétérinaires, Laboratoires vétérinaires départementaux, abattoirs, Météorologie nationale. Cette collecte d'information auprès d'organismes tiers est indispensable pour des informations qui ne sont pas disponibles au sein des élevages. Cela permet également d'éviter les activités de collecte en "double", donc de limiter le travail de l'enquêteur à des relevés d'informations non formalisées par ailleurs.

L'origine multiple des informations nécessite de résoudre les problèmes liés à l'acquisition et à la mise en correspondance d'informations de sources différentes. Ces problèmes sont par exemple :

- La propriété des données (récupération des données du Contrôle laitier par exemple),
- Les aspects juridiques, régis notamment par la loi Informatique et libertés (cf. encart 2),
- L'identification des individus (correspondance entre les N° d'identification EDE collectés dans l'enquête et ceux contenus dans les fichiers du Contrôle Laitier),
- Le support de l'information : documents parfois mal adaptés à la saisie, difficulté d'interface informatique,
- La contradiction entre informations de sources différentes [Faye et al., 1994] pour lesquelles il faut déterminer une procédure de choix (par exemple, date du dernier contrôle laitier postérieure à la date du tarissement fournie par l'éleveur).

Encart 2**INFORMATIQUE ET LIBERTES**

Les E.E.A. relèvent de la loi "informatique et libertés" au sens où "les traitements automatisés d'informations nominatives" doivent, préalablement à leur mise en oeuvre, faire l'objet d'une déclaration auprès de la Commission nationale de l'informatique et des libertés (C.N.I.L.). En effet, un traitement nominatif peut être soit directement nominatif, soit indirectement nominatif : par directement nominatif, on entend tout traitement qui utilise les noms des personnes physiques ; par indirectement nominatif, on entend tout traitement qui peut permettre d'identifier une personne physique sans que son nom apparaisse en clair (ex : n° matricule ou code, n° de téléphone, recoupement de plusieurs renseignements informatisés non nominatifs).

La gestion d'une E.E.A. implique généralement le recueil des noms et coordonnées de personnes physiques (éleveur, enquêteur, technicien...), ce qui est considéré comme des traitements directement nominatifs (courriers aux partenaires de l'enquête, retours d'information). Par ailleurs, même dans le cas où les individus statistiques sont anonymes (enquête pour un tiers par exemple), le recoupement des informations recueillies (N° I.N.S.E.E. de commune, identification des animaux, éléments socio-économiques des élevages...) peut permettre de "retrouver" la personne physique ; cela correspond à une possibilité de traitements indirectement nominatifs. Dans la majorité des cas, il faut donc faire une déclaration d'un traitement automatisé d'informations nominatives.

Il est à noter que les traitements mis en oeuvre dans les E.E.A. sont classés, au regard de la loi, dans les types de traitement "Statistiques" ou "Clients" et qu'ils font l'objet de normes et de déclarations simplifiées.

En pratique :

- déclaration simplifiée avant la mise en place de l'enquête (1 seul formulaire, 15 rubriques à remplir),
- attente du récépissé de déclaration (moins de 15 jours en général).
- séparation physique des données : un fichier avec les informations directement nominatives (nom, adresse, téléphone) et les identifiants anonymes (code numérique) qui seront utilisés dans les autres fichiers ; ainsi, la suppression des relations entre les fichiers (en désactivant les liens dans la base de données) permet une séparation physique des données nominatives et non nominatives.
- quelques lignes sur les questionnaires (qui mettent en jeu des informations nominatives) où on indique la destination des informations, et on précise que les personnes physiques ont le droit de consultation et de rectification des informations qui les concernent.

La référence : la loi n° 78-17 du 6 janvier 1978

Les modes d'emploi :

- le formulaire C.E.R.F.A. n° 99001 et la notice explicative de la loi (à retirer auprès des Préfectures, des Chambres de commerce et d'industrie, ou de la C.N.I.L.).
- la brochure 1473 intitulée "Informatique et libertés" (à demander au Journal officiel)

Les adresses :

- pour les déclarations : Commission nationale de l'informatique et des libertés, 21 rue Saint-Guillaume, 75007 Paris Tél : (1) 45 44 40 65
- pour les textes et documents : Journal officiel, 26 rue Desaix, 75015 Paris

C - COMPLEXITE**1 - NIVEAUX D'INFORMATION**

L'information mise en jeu dans les E.E.A. correspond à différents individus (appelés "entités" en langage informatique), l'exploitation, l'éleveur, l'animal, etc. On peut décrire chacun de ces individus par des propriétés, des variables (pour une vache, son N° d'identification E.D.E., sa race, sa date de naissance...).

D'autre part, l'information est structurée en niveaux : l'animal, le lot, le bâtiment, l'élevage (cf. figure 1). Cette structuration peut être plus fine et utiliser plusieurs niveaux pour le même individu (résultats sérologiques, pathologie, appréciation de l'état d'engraissement, mesures de production laitière...). Par exemple, un fichier contenant les pesées hebdomadaires de production laitière des vaches est structuré au niveau de la pesée [Calavas, 1992 a].

L'existence de ces différents niveaux d'information a deux grandes conséquences. D'une part, cela rend nécessaire la modélisation des données car une structuration de l'information "individu - propriété" devient insuffisante (pour donner une image, on peut dire que l'on n'arrive pas à définir l'information avec un seul tableau où les lignes seraient les individus et les colonnes représenteraient les propriétés caractérisant ces individus). Il faut définir des relations entre les différentes entités et les différents niveaux par exemple avec le modèle "entité - relation" [Tardieu et al., 1983].

D'autre part, il faut prévoir l'identification des individus présents à chaque niveau d'information (il faut pouvoir identifier sans ambiguïté le cycle de production, l'animal, le lot d'animaux, le bâtiment, l'élevage...) et prévoir les liens qui permettent de faire correspondre ces identifiants entre les niveaux d'information (par exemple, retrouver toutes les pesées de production laitière pour calculer la production annuelle d'une vache).

2 - GESTION DU TEMPS

Les E.E.A. nécessitent de gérer de nombreuses données temporelles (dates, heures...). Ces données sont nécessaires pour vérifier la chronologie des événements (une date de métrite est postérieure à la date du vêlage correspondant), quantifier des informations (critères de fécondité définis par des intervalles de temps) ou déterminer la période de survenue d'un événement (début d'un cycle de production, apparition d'une pathologie)... Les données temporelles sont des données sensibles car difficiles à mémoriser et soumises à de nombreux risques d'erreur [Schukken et al., 1989]. Par ailleurs, elles sont souvent furtives.

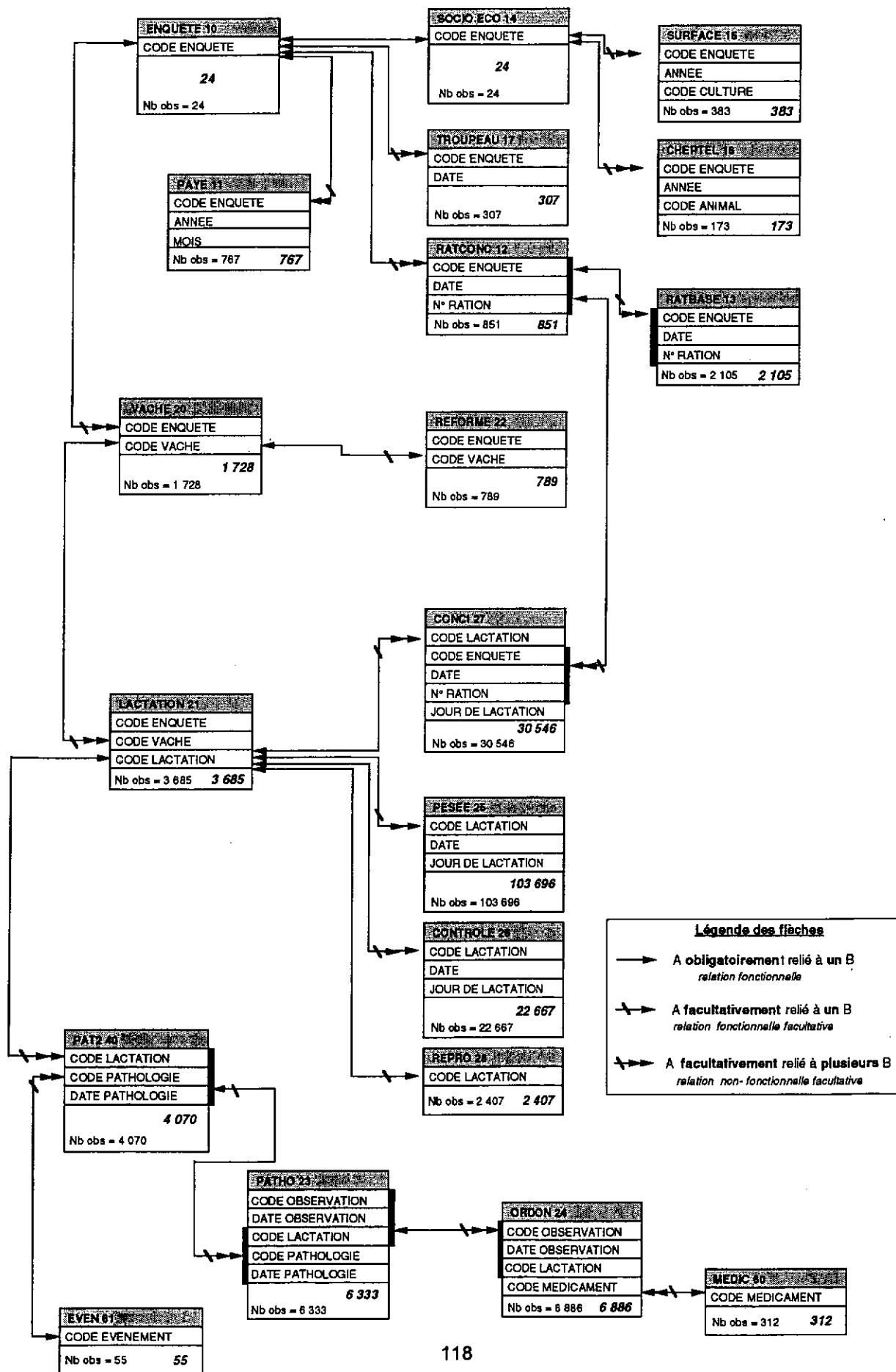
En plus de la difficulté de collecte de ces informations, la notion de temps doit être gérée au niveau de l'enquête car le temps s'accompagne de l'évolution des effectifs (entrées et sorties d'individus de l'enquête). La ré-observation des individus (animal qui commence un nouveau cycle de production) nécessite une gestion précise, car c'est souvent un seul événement (la mise-bas) qui déclenche la clôture d'un enregistrement (fin de la lactation n) et le début d'un autre (début de la lactation $n+1$).

Cette complexité doit être prise en compte très tôt dans la conception de l'enquête car elle a des répercussions sur l'identification des individus (animaux, élevages), l'organisation des questionnaires, la stratégie de collecte et d'analyse des données (faisabilité de la collecte au niveau de l'animal, analyse de l'information au niveau de l'animal ou au niveau de l'élevage...).

Les volumes d'information dans les E.E.A. (10^5 à 10^8 données) rendent incontournables l'utilisation de l'informatique (cf. encart 3).

Mais c'est surtout la nature, la diversité, la complexité de l'information recueillie dans les E.E.A. qui rendent nécessaire de penser, avant que l'enquête ne commence sur le terrain, à l'organisation du recueil de l'information (conception de l'enquête, conception de la base de données), aux étapes entre le recueil de l'information et le chargement de la base de données (acheminement des données, vérification, saisie), et enfin à l'utilisation de la base de données (stratégies d'exploitation et documentation des données).

Figure 1 : Schéma organisationnel du référentiel de données de l'enquête « Coûts des pathologies en élevage bovin laitier » (Centre d'écopathologie animale)



Encart 3

QUEL SYSTEME INFORMATIQUE ?

Il y a quelques années, une E.E.A. nécessitait l'utilisation de "gros systèmes informatiques", c'est-à-dire des ordinateurs de type mini, multi-tâches, seuls capables alors d'offrir la puissance de calcul nécessaire aux traitements (manipulations de tables, modélisation, analyse des données...), et de gérer des gros volumes (mémoire vive importante, grande capacité de stockage).

Aujourd'hui on peut gérer, du point de vue informatique, une enquête de taille raisonnable (5×10^5 données) avec un matériel peu coûteux (micro-ordinateur compatible PC type 486, 4 Mo RAM, 200 Mo disque). Cependant, les applications dites spécialisées (Base de données, calculs scientifiques) requièrent en général un supplément de puissance (préférer 8 Mo de mémoire vive, et un processeur à haute cadence 33 ou 66 Mhz), notamment si on travaille sous interface graphique (Windows).

Les logiciels utilisés dans les E.E.A. sont de nature diverse : S.G.B.D., traitement de texte, P.A.O., cartographie, tableur, calcul statistique. Néanmoins, outre les logiciels statistiques indispensables à l'exploitation des données recueillies, deux autres types de logiciels sont vraiment indispensables : un traitement de texte pour les échanges avec les partenaires (mailing, envoi des conventions, courriers de suivi de l'enquête, relances, retours d'information...) et un système de gestion de base de données (S.G.B.D.). De nombreux logiciels sont disponibles sur le marché. En ce qui concerne le traitement de texte, il faut veiller aux possibilités d'interface avec les fichiers provenant de la base de données.

Pour les S.G.B.D., le choix est plus large parmi les produits que l'on peut classer en trois grandes familles :

- Les S.G.B.D. compatibles avec Dbase (Dbase, Foxpro, Clipper...)
- Les S.G.B.D. "grand public" (Paradox, Access...) : interface performante, limites du langage de développement,
- Les S.G.B.D. spécialisés (Oracle, Ingres, Progress, Sybase, Dataflex, RDB, DB2...) : "moteurs" très performants (SQL, sécurité des données, outils de développement), disponibles sur plusieurs systèmes d'exploitation, malgré une interface parfois lourde (notamment sur gros systèmes).

Le choix d'un S.G.B.D. est difficile et on ne peut détailler ici tous les critères de choix. Toutefois, compte tenu du contexte d'utilisation dans les E.E.A., on peut retenir la rapidité de création de la base de données (par exemple, le temps de création des interfaces, masques de saisie et menus ne doit pas obérer le développement des procédures de contrôle de la qualité des informations), la facilité d'interface (notamment avec les logiciels statistiques), la fonctionnalité de l'outil d'interrogation de la base (puissance du langage, SQL par exemple, ergonomie de l'outil de requête...). On peut être amené à utiliser également des logiciels de publication (P.A.O.) pour réaliser les questionnaires d'enquête, et des outils de présentation (Pré.A.O.) pour la mise en forme des sorties graphiques. Enfin, les échanges de données avec les partenaires peuvent être envisagés par messagerie électronique (via des réseaux nationaux et internationaux) ce qui nécessite un modem pour accéder au réseau téléphonique ou un accès à un réseau Ethernet.

Pour des enquêtes plus lourdes ou pour un plus grand confort d'utilisation, on utilisera les mêmes logiciels sur des plates-formes plus performantes (Unix, Vax/VMS, MVS...) installées sur des ordinateurs de type mini ou des réseaux. Sur de tels systèmes, une organisation client-serveur permet l'économie et le partage non seulement des ressources (disques, imprimantes, traceurs, système de sauvegarde), mais aussi des applications (S.G.B.D., logiciels statistiques, outils bureautiques...) et des données (base de données, fichier des partenaires...). En revanche, tous les logiciels existants pour la micro-informatique ne sont pas disponibles ou sont coûteux sur certaines plates-formes ; on peut ainsi être amené à construire un réseau informatique hétérogène composé de différents types d'ordinateurs (PC, Macintosh, mini, station de travail...) en utilisant le meilleur de chaque monde (puissance et sécurité de traitement des minis, graphisme des stations de travail, communication et échanges de données sur les PC).

II - ORGANISATION DU RECUEIL DE L'INFORMATION

A - CONCEPTION DE L'ENQUETE

Le recueil des données * dans les E.E.A. est en général confié à des intervenants habituels de l'élevage (vétérinaires, techniciens...). Ce choix présente des avantages (connaissance technique, relation privilégiée avec les éleveurs, motivation) mais introduit des contraintes liées à la diversité de ces enquêteurs (niveau de connaissance hétérogène, disciplines diverses), au fait que le travail d'enquête est un travail supplémentaire et à l'implication relationnelle forte entre l'organisme de recherche et les enquêteurs d'une part, entre les enquêteurs et les éleveurs d'autre part [Calavas et al., 1994].

Des moyens méthodologiques ont été développés pour remédier à ces contraintes : la conception de l'enquête en groupe de travail pluridisciplinaire, la préenquête, la formation des enquêteurs, le suivi pendant la réalisation de l'enquête [Martin et al., 1987].

Au stade de la conception de l'enquête, la fonction du groupe de travail est de rechercher, d'identifier, de valider et de formaliser les hypothèses de recherche (facteurs de risque par exemple) et de définir le protocole d'enquête. Le groupe de travail, de par sa connaissance technique globale et ses connaissances des contraintes de la filière, est une garantie de la faisabilité de l'enquête qu'il conçoit. Il est important que certains membres aient une expérience d'enquêteurs, car en se "voyant" déjà dans le rôle d'enquêteur quand ils conçoivent l'enquête, ils réfléchissent d'autant plus à la faisabilité, notamment en ce qui concerne le recueil des données.

La préenquête rend possible la mise au point d'instruments et de méthodes de mesure et de notation qui permettent de formaliser l'information. Il s'agit par exemple de concevoir des outils pratiques de mesure (appareil permettant de mesurer la circonférence du carpe et du métacarpe des chèvres pour la caractérisation clinique du "gros genou"

[Monicat, 1988]), des outils de notation semi-quantitatifs (appréciation de l'état de propreté des vaches [Faye et Barnouin, 1985]), d'harmoniser des protocoles d'analyses biologiques entre différents laboratoires [Peretz, 1992].

Une des principales fonctions de la préenquête est de valider le projet de protocole (vérification du devis de temps de recueil, modalités d'acheminement des données) et d'optimiser les supports de recueil des données en terme d'ergonomie (questionnaires structurés chronologiquement en fonction de l'organisation des visites, explications et éléments de codage des réponses fournies simultanément aux questions et aux actes prévus) [Rumeau-Rouquette et al., 1993] et en terme de lisibilité (polices de caractère différentes pour les questions et les commentaires, bacs de codage informatique regroupés dans une colonne).

La formation des enquêteurs est nécessaire en raison de leur profession et discipline diverse, de leur niveau d'implication et de connaissance différent vis-à-vis de l'enquête. Les sessions de formation ont pour but, outre l'exposé des objectifs de l'étude et de son protocole, la standardisation du recueil des données et une sensibilisation aux contraintes de celui-ci (utilisation des questionnaires, problème des "données manquantes", importance de l'identification des individus).

B - CONCEPTION DE LA BASE DE DONNEES

1 - MODELISATION DES DONNEES

La réalisation technique d'une base de données doit s'appuyer sur une conception par étape, pour laquelle des méthodes dédiées aux systèmes d'information comme MERISE [Tardieu et al., 1983] fournissent une démarche et des outils.

* L'usage a consacré l'utilisation du terme « recueil des données » pour ce qui concerne l'acquisition de l'information

Après avoir défini l'"univers des données" idoine, il s'agit de construire un modèle sémantique de l'information, à un niveau purement conceptuel qui ne préjuge donc pas de choix techniques. La modélisation conceptuelle est particulièrement importante pour les performances de la future base de données : redondance réduite, entrée et extraction des données aisées, adaptation à des applications diverses, évolution. Elle est basée dans MERISE sur une approche voisine du schéma "entité-relation" de Chen [1976], qui identifie dans l'ensemble des types de données élémentaires et leurs liens, des entités (vache, élevage, événement...) et des relations (appartenance d'une vache à un élevage). L'optimisation du modèle conceptuel, visant les performances évoquées plus haut, peut être assurée par l'emploi de techniques de Cood [1972, 1974 ; figure 2]. Elle renvoie aussi à la prise en compte de l'utilisation future des données. La figure 3 présente un modèle conceptuel où l'entité "rang de lactation" témoigne de la prédilection pour la "vache-lactation" comme unité d'observation des phénomènes étudiés.

L'étape suivante est une traduction triviale du modèle conceptuel tenant compte de choix techniques (construction d'un schéma relationnel par exemple).

Le rôle clef de la représentation conceptuelle, qui est non seulement un préalable à la base de données mais aussi un outil de dialogue entre les partenaires de l'enquête, a conduit Lescourret et al. [1993] à décrire des modèles de données utiles aux enquêtes épidémiologiques sur les vaches laitières.

2 - CONCEPTION SIMULTANÉE A CELLE DE L'ENQUÊTE

La conception et la création de la base de données sont des étapes qui déterminent les performances du circuit de l'information. Le travail de conception, lorsqu'il est mené de façon simultanée à la conception de l'enquête [Sulpice et al., 1994], limite les "pertes de charge en ligne" entre conception de l'enquête et conception de la base de données : cela contribue à la réduction des délais entre la fin du recueil des données et le début du traitement statistique.

Par ailleurs, cela permet de concevoir un circuit d'information où la base de données est

opérationnelle dès le début du recueil des données, et voit donc son rôle renforcé (cf. figure 4 et encart 4). La base de données n'est plus seulement le réceptacle des données recueillies, elle assure alors 4 fonctions complémentaires, qui sont l'aide à la conception de l'enquête, la gestion de l'enquête, le contrôle de la qualité des données et la préparation des données pour l'analyse statistique.

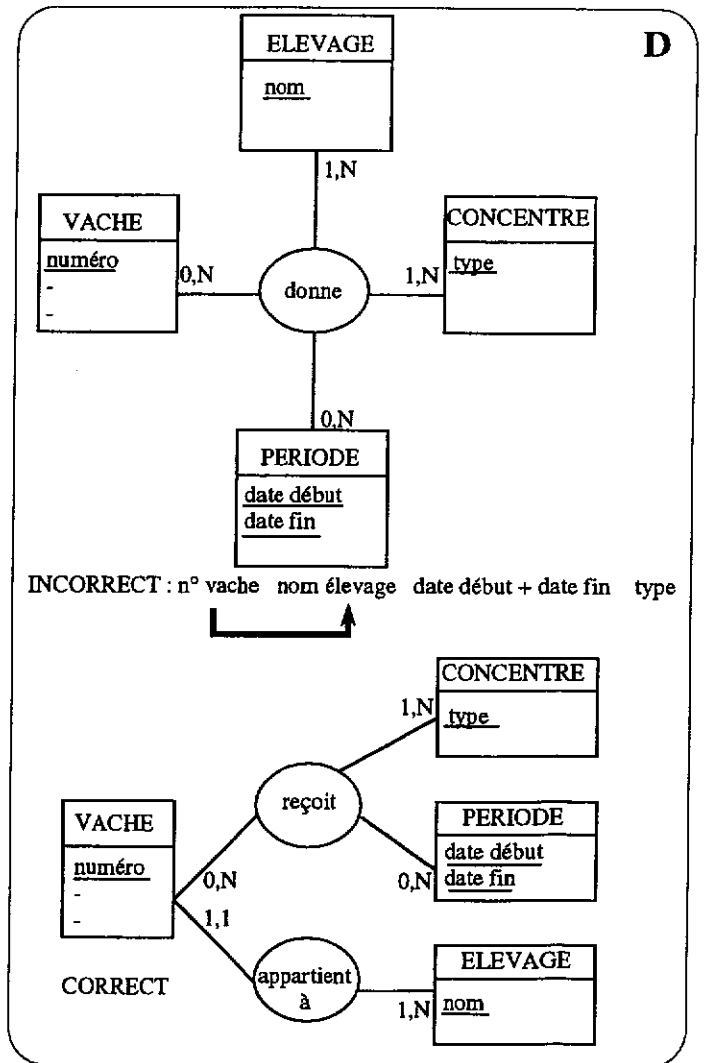
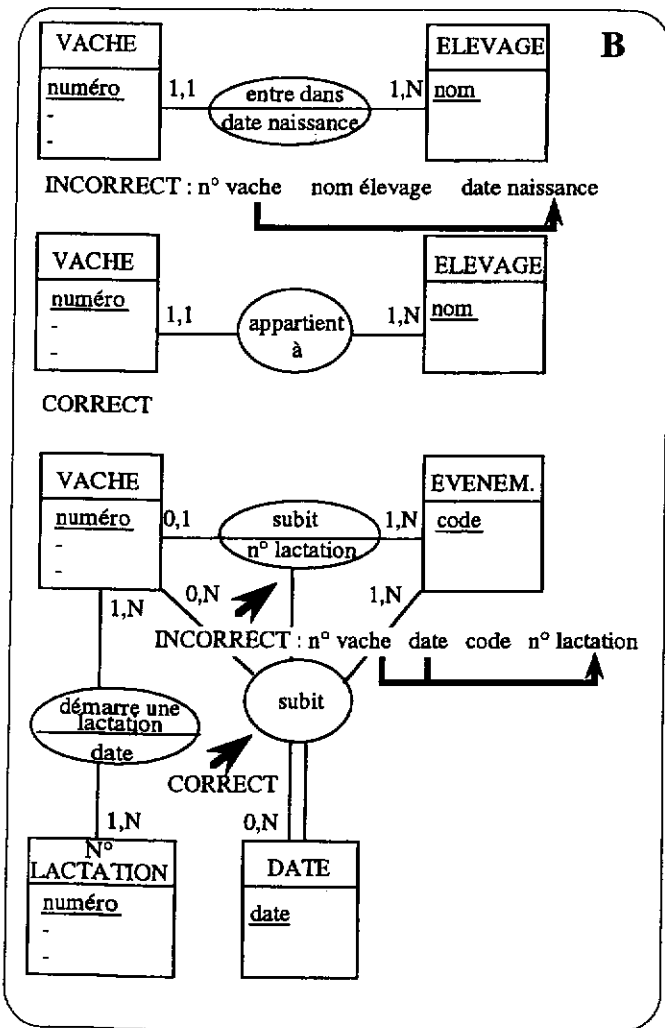
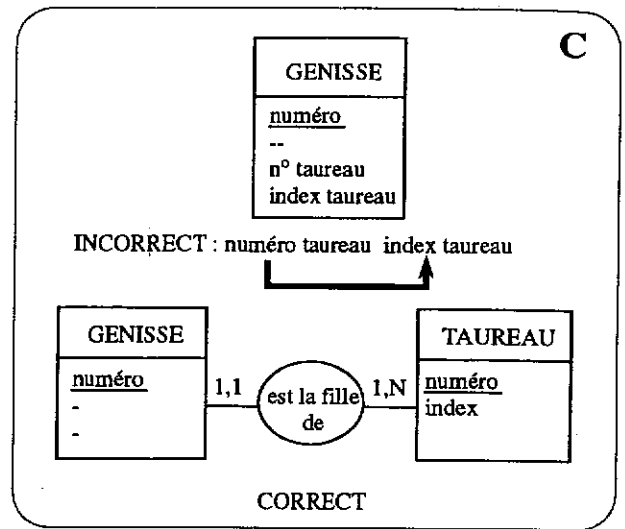
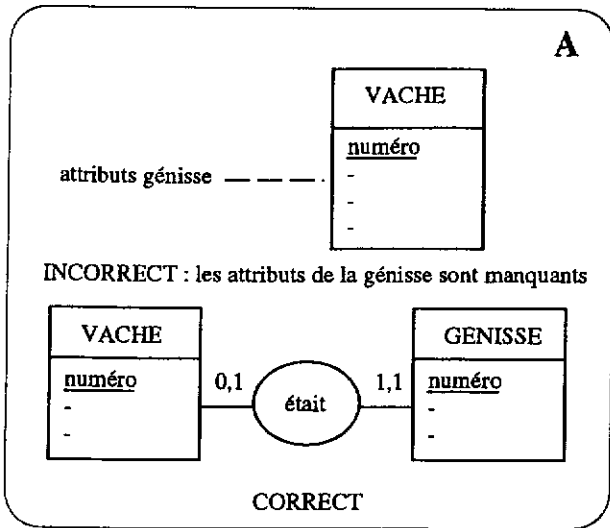
3 - APPORTS DE L'INFORMATIQUE A L'ENQUÊTE

Ce travail simultané de conception de l'enquête et de la base de données permet d'intégrer des contraintes informatiques dans la conception de l'enquête. Il permet également d'enrichir l'enquête en intégrant l'informatique, au sens de science de l'information, dans la démarche pluridisciplinaire de conception.

Parmi les apports de l'informatique à la conception des enquêtes, on peut citer comme les plus importants :

- L'identification des individus : une solution simple est, par exemple, de choisir un identifiant animal incluant l'identifiant élevage. Ce type de codage (dit "codage hiérarchisé" ou "à niveaux") possède une assez bonne lisibilité pour l'éleveur et l'enquêteur (dans un élevage donné, on gère les individus uniquement avec l'identifiant animal) et permet d'un point de vue strictement informatique de faciliter les liens entre les fichiers de "niveau élevage" et ceux de "niveau animal".
- Le codage des informations : un "oeil" informatique s'attachera à vérifier l'exhaustivité du codage des questions fermées, notamment en s'assurant que le codage des "données manquantes" et des "données sans objet" est prévu.
- La mise en place de variables de statut des individus : par exemple, dans une enquête sur les mammites de brebis allaitantes [Calavas, 1992 b), une variable "statut des brebis lors de la seconde visite" (présente, réformée, vide, morte, perdue de vue) facilite le suivi précis des individus et donc permet de diminuer les "perdus de vue" qui sont les "bêtes noires" des E.E.A.

Figure 2 : Contrôle des modèles à l'aide des techniques de Cood
 Première forme normale (A), seconde (B), troisième (C) et forme normale de Boyce-Cood (D)
 (Laboratoire d'écopathologie-INRA Theix)



- objet
- relation ne portant pas d'attribut
- relation portant des attributs
- attributs
- xxx identifiants
- 0,N cardinalité
- dépendance fonctionnelle

Encart 4

GESTION DE L'INFORMATION DANS 2 E.E.A. : DIFFERENCES ET SIMILITUDES

Le premier système d'information (cf. figure 5) est celui de l'enquête écopathologique Bretagne du Laboratoire d'écopathologie de l'I.N.R.A. de Theix [Faye et al., 1989]. Cette enquête est dédiée à la mise en évidence des facteurs de risque de diverses pathologies du *peri partum* chez la vache laitière, à la recherche des associations entre ces pathologies et à celle de leurs liens avec les performances (qualité du lait...). Compte tenu de la variété de ses objectifs, ce projet se situe dans un cadre de recherche exploratoire.

Les volumes d'information mis en jeu, la multiplicité des sources d'information, et la durée d'enquête ont impliqué la mise en place d'un système d'information temporaire, dédié à la gestion du recueil des données (saisie, contrôle de cohérence, stockage de données déjà informatisées...). La structuration des informations se fait par item, et les données sont stockées dans une collection de fichiers indépendants. Après leur validation, les données sont transférées dans un système d'information définitif.

La base de données relationnelle a été conçue en privilégiant l'indépendance des données vis-à-vis des analyses statistiques, et doit supporter un langage d'extraction (SQL) permettant aux chercheurs de développer facilement toutes les stratégies d'analyse correspondant aux objectifs variés de l'enquête.

Le second système d'information (cf. figure 4) concerne une enquête d'écopathologie sur les mammites des brebis allaitantes (Calavas, 1992 b) ; cette enquête mise en place par le Centre d'écopathologie animale est ciblée sur un objectif : la mise en évidence des facteurs de risque des mammites des brebis allaitantes. Le recueil des données a été réalisé dans 78 exploitations avec une durée d'observation de 3 à 4 mois par élevage.

Ce système est orienté davantage vers l'informatisation de la gestion des données d'enquête permettant l'obtention rapide de la base de données. La stratégie d'analyse statistique a été formalisée lors de la conception de l'étude. Le recueil des données est donc adapté aux hypothèses de recherche, et on peut dire schématiquement que les entités (ici "Description des brebis", "Pratiques de l'éleveur lors du sevrage des agneaux", "Sérologies", etc.) ont été traduites en dossiers, puis en questionnaires.

La base de données a deux objectifs, assurer la gestion de l'enquête (gestion des données, des acteurs, des flux d'information, contrôle de la qualité) et être un serveur d'information pour les analyses statistiques. La base assure donc l'interface avec les logiciels statistiques auxquels elle fournit des données documentées (données et dictionnaires des données).

Ces deux enquêtes sont globalement différentes notamment en terme d'objectifs de recherche, recherche exploratoire pour l'une, recherche de facteurs de risque d'une pathologie définie pour l'autre. Toutefois, la comparaison des deux systèmes d'information, au delà des particularités liées aux objectifs, montre que la gestion de l'information a été abordée de manière similaire, à savoir :

- la gestion de l'information a été prévue dès la conception de l'enquête,
- le contrôle de la qualité des données fait partie intégrante du système d'information avec notamment la détection la plus précoce possible des erreurs et le retour à la source en cas d'erreur,
- une dynamique d'échange d'information avec les acteurs de l'enquête (retour d'information aux partenaires) a été mise en place pour faciliter le recueil de l'information,
- l'organisation de l'information a nécessité la mise en place d'une base de données relationnelle, qui accorde une grande place à la documentation des données (dictionnaire), et à l'interface avec les utilisateurs (langage d'extraction des données, passage des données vers les logiciels statistiques

III - GESTION DE L'INFORMATION PENDANT L'ENQUETE : DU RECUEIL DES DONNEES A LA BASE DE DONNEES

A - ACHEMINEMENT DE L'INFORMATION

Le transfert des questionnaires d'enquête entre le lieu de collecte et le lieu de traitement informatique de l'information (par acheminement postal dans la plupart des cas) peut se faire avec un rythme différent selon le protocole (durée de l'enquête, nombre de visites, origine des informations). D'une manière générale, il faut limiter le délai entre le recueil des données dans l'élevage et leur réception sur le lieu de traitement informatique. L'acheminement des questionnaires peut se faire en fin d'enquête si la durée de celle-ci est courte (inférieure à 2 mois) ; pour des enquêtes plus longues, il faut privilégier l'acheminement des questionnaires après chaque phase de recueil des données (chaque visite de troupeau par exemple).

L'acheminement des questionnaires en cours d'enquête permet d'instaurer une dynamique d'échange : on peut, par exemple, réaliser l'identification des animaux lors d'un premier passage dans l'élevage, prévoir l'expédition des questionnaires au centre de recherche où les informations sont informatisées, ce qui permet, d'une part, de vérifier les données recueillies, et, d'autre part, de produire des outils (inventaire de troupeaux, étiquettes pour identifier des prélèvements de sang, etc.) facilitant le travail des enquêteurs lors de la seconde visite (cf. figure 4) [Calavas, 1992 b ; Ducrot, 1992].

Une logistique doit être mise en place, qui vise à limiter les pertes d'information pendant cette phase : colis résistant, affranchissement suffisant, étiquettes pré-imprimées.

B - CHARGEMENT DE LA BASE DE DONNEES

Le chargement de la base de données ne se limite pas à la saisie informatique (cf. figure 5). C'est un des moments clé du contrôle de la qualité des informations. On peut en préciser trois grands aspects : les étapes d'acquisition des données (relecture, correction, saisie en continu), les contrôles de cohérence, le retour à la source des informations [Pons, 1989].

C - ACQUISITION DES DONNEES

Les étapes de relecture, de correction, et de saisie informatique des documents peuvent être réalisées en lot à la fin du recueil des données, ou de façon continue. Dès que la durée de recueil des données dépasse 2 mois, on peut conseiller de travailler de façon continue, pour réduire les délais et pour garder la possibilité de retourner à la source des informations.

L'acquisition automatique des informations par lecture optique (reconnaissance de caractères, codes barres) est rarement utilisée car elle est lourde à mettre en oeuvre et n'est économiquement justifiée qu'avec des volumes d'information beaucoup plus importants que ceux que l'on a dans les E.E.A. En revanche, la collecte informatique des données (i.e. l'utilisation de micro-ordinateurs portables) sur le site est plus fréquemment mise en oeuvre, notamment dans les abattoirs [Brunet, 1985] et les laboratoires. Cette technique, très intéressante en ce qui concerne la fiabilité des données, est parfois lourde à appliquer dans les élevages, en raison de l'investissement humain et du temps qu'implique son utilisation [Sulpice, 1992].

Figure 4 : Déroulement de l'enquête « Mammmites des brebis allaitantes » et articulation avec la base de données (Centre d'écopathologie animale)

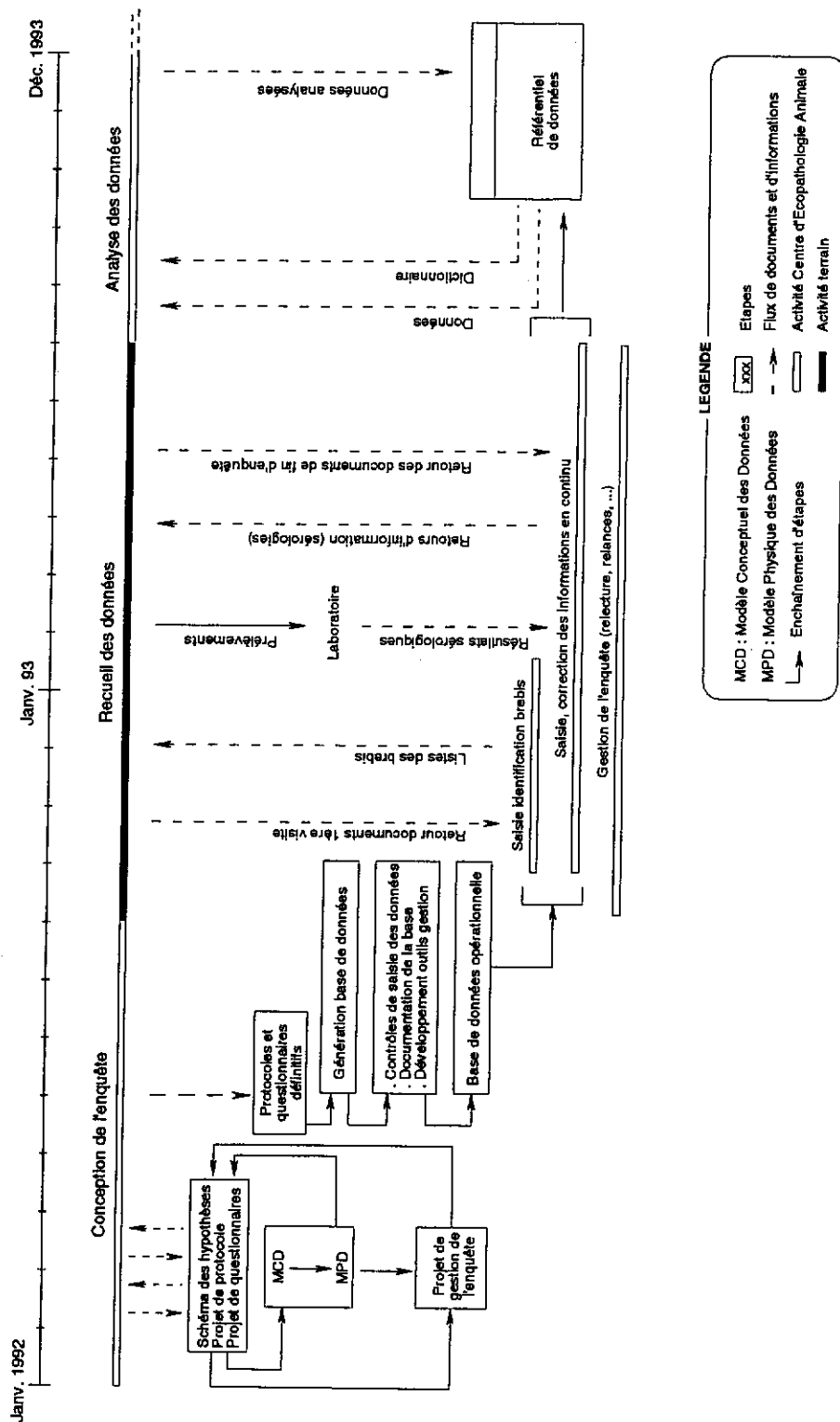
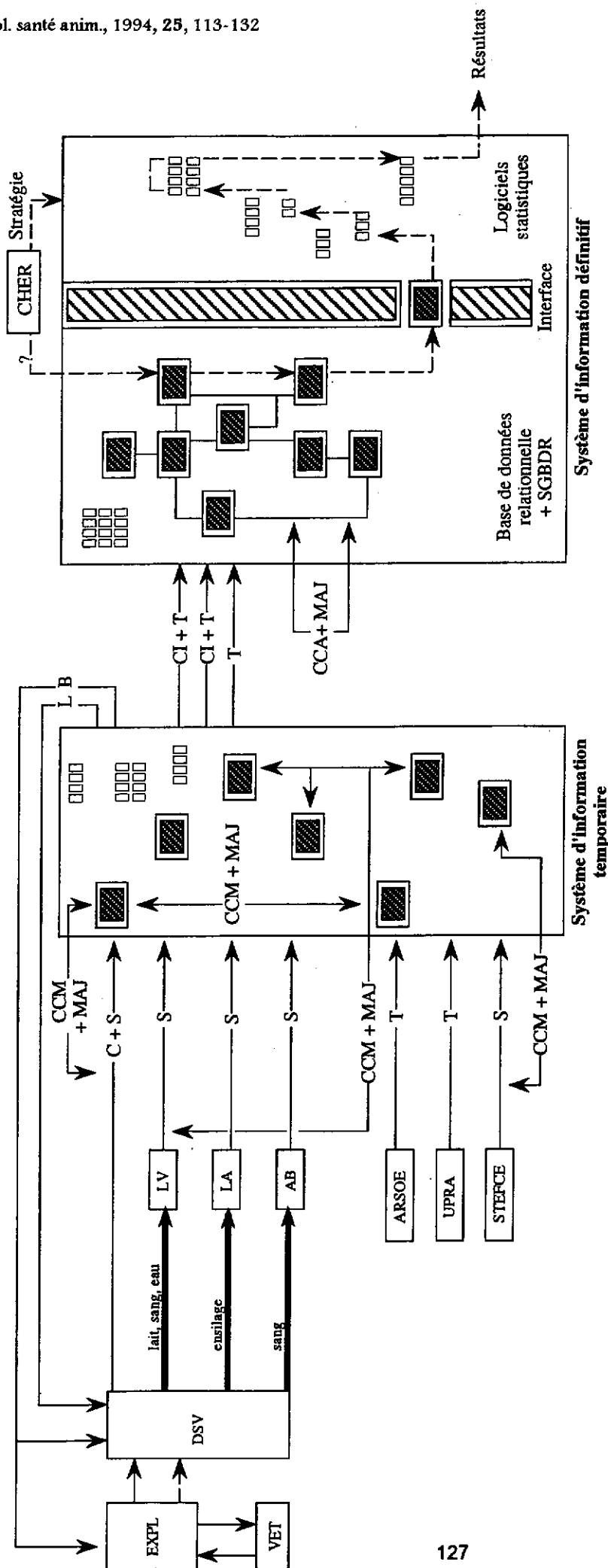


Figure 5 : Diagramme de flux de l'enquête d'écopathologie Bretagne (Laboratoire d'écopathologie - INRA Theix)



	Acteurs		Fichiers, tables ou tableaux de données	CCM	Contrôle de cohérence manuel
	EXPL : Exploitations		Applications informatiques	CCA	Contrôle de cohérence automatisé
	VET : Vétérinaires praticiens		Flux d'information	MAJ	Mise à jour
	DSV : Direction des Services Vétérinaires (techniciens enquêteurs)		Flux de matière	L	Création de listages (support d'informations)
	LV : Laboratoire vétérinaire départemental (analyses)		Chaine de traitement	B	Création de bilans
	LA : Laboratoire départemental d'analyses alimentaires		C + S Codage + Saisie	CI	Création contrôlée de l'image des tables
	AB : Atelier biochimie du laboratoire d'écopathologie (analyses)		S Saisie contrôlée	?	Requêtes (création de tableaux de données)
	ARSOE : Association Régionale de Service aux Organismes d'Élevage (Bretagne)		T Transfert (de données informatisées)	SGBDR	Système de gestion de bases de données relationnelles
	UPRA : Union de Sélection et de Promotion de Race				
	STEFCE : Service Technique des Facteurs Climatiques de l'Environnement				
	CHER : Chercheurs				

D - CONTROLES DE COHERENCE

Le contrôle des données est une étape lourde et souvent longue, c'est pour cela qu'il convient de préciser et de prévoir le mode d'application de chacun de ces contrôles, ceux-ci devant être adaptés à chaque type d'erreur [Faye et al., 1994] :

- Structure des données (type et format) : lors de la saisie,
- Vraisemblance (minimum, maximum, possibilité de "données manquantes" ou "sans objet") : lors de la saisie,
- Cohérence des valeurs entre les différentes variables d'un individu (cohérence interne) : lors de la saisie (questions conditionnelles et données sans objet) et/ou par lot (phase de vérification des données),
- Cohérence avec d'autres tables (cohérence externe) : par lot sauf pour des données très sensibles (identifiants des individus, variables à expliquer dans une étude).

Le codage des données et l'organisation des questionnaires ne sont pas sans influence sur la mise en oeuvre, plus ou moins aisée des contrôles. Une même information, par exemple la race d'une vache sera plus facile à vérifier si on choisit un codage numérique continu (1 pour Prim'Holstein, 2 pour Montbéliarde, 3 pour Charolaise...) plutôt qu'un codage énuméré (66 pour Prim'Holstein, 46 pour Montbéliarde, 38 pour Charolaise). Il est en effet plus facile de faire un contrôle par une fourchette de valeurs (minimum, maximum) qu'un contrôle par des valeurs énumérées.

L'organisation des questionnaires facilite quant à elle la réalisation des contrôles de cohérence interne et externe. C'est le cas des questions conditionnelles (questions qui ne s'appliquent qu'en fonction de la réponse à une autre

question). Lorsque ces questions conditionnelles n'ont pas à être posées, elles génèrent des "données sans objet". On veillera donc à un enchaînement adapté entre les questions posées systématiquement et les questions conditionnelles, ce qui permettra de gagner du temps lors du renseignement du questionnaire ("si le bâtiment est une stabulation libre, passer à la page suivante"), et de repérer facilement les "données sans objet" lors de la saisie.

**E - RETOUR A LA SOURCE
DE L'INFORMATION**

Le retour à la source de l'information en cas de "données manquantes", ou en cas d'erreur sur les données [Sulpice, 1993], doit être prévu à l'avance afin qu'il soit réalisé dans de bonnes conditions. Pour cela, la collecte de compléments d'information doit être humainement réalisable et mise en oeuvre alors que le recueil des données est encore en cours pour éviter toute démobilité des partenaires, inévitable après la fin de l'enquête. D'un point de vue technique, et dans un souci d'économie de temps et de moyens, on peut si nécessaire envisager une automatisation des relances (formulaire pré-établi à compléter manuellement avec les questions précises aux partenaires), la gestion de ces relances étant quant à elle toujours nécessaire (gestion des départs de relance, des arrivées, mises à jour de la base de données).

Le retour à la source de l'information apparaît comme une solution de secours, de rattrapage permettant d'améliorer l'exhaustivité et la fiabilité des données dans les enquêtes. Du point de vue de l'efficacité, il faut réserver les relances aux informations pérennes (cf. encart 1) afin de ne récupérer que des informations fiables ; il est en effet inutile d'essayer de récupérer des informations furtives non recueillies en raison d'une insuffisance du protocole ou de la logistique de l'enquête.

IV - REGLES ET MODALITES D'UTILISATION DE LA BASE DE DONNEES

A - OBJECTIFS ET REGLES DE LA BASE DE DONNEES

Trois éléments clés permettent de maintenir l'intégrité des données et d'atteindre les objectifs de l'enquête : la protection des données, la facilité d'accès aux données, la souplesse d'utilisation.

1 - PROTECTION DES DONNEES

On pourra par exemple interdire la mise à jour (modification, ajout, suppression d'enregistrement) de la base de données après les dernières vérifications, en autorisant uniquement l'accès en lecture des données [Martin, 1977].

2 - ACCESSIBILITE

Il faut pouvoir accéder aux données de deux manières : par un accès direct à chaque table pour la consultation des informations et par des extractions mettant en jeu plusieurs tables en vue des analyses statistiques. La facilité et la rapidité d'accès dépendent essentiellement de la modélisation et de la documentation de la base de données (dictionnaire permettant de situer telle information dans l'ensemble des tables) et secondairement de la puissance du logiciel (procédures d'extraction de données plus ou moins performantes).

3 - SOUPLESSE D'UTILISATION

La base de données joue le rôle de serveur d'information, et il ne doit pas y avoir de contraintes d'accès aux informations. Pour une base de données vouée à la recherche exploratoire, une indépendance entre la structure des données et les analyses à effectuer est recherchée [Lescourret et al., 1992].

D'autre part, la base de données doit pouvoir être enrichie avec des informations produites par l'analyse statistique. Cela peut être réalisé par exemple en structurant la base de données en deux couches : une couche "fermée" (interdite à la mise à jour) contenant les données recueillies

dans l'enquête, et une couche "ouverte" aux importations de données calculées et de résultats.

B - UTILISATION DE LA BASE DE DONNEES

La base de données contient toutes les informations de l'enquête, données recueillies, dictionnaire des données. Elle doit être accompagnée des questionnaires référencés, des masques de saisie et de son mode d'emploi. Cet ensemble permet à un non-informaticien, et/ou à quelqu'un n'ayant participé ni à la conception de l'enquête ni au recueil des données de pouvoir utiliser les informations de l'enquête : c'est la notion de référentiel de données [Sulpice et al., 1994].

1 - DOCUMENTATION

L'objectif du dictionnaire des données est d'établir la correspondance entre le signifiant et le signifié (cf. encart 5.). Cette correspondance est établie lors de la création de la base de données, soit en utilisant le dictionnaire des données propre au système de gestion de base de données (S.G.B.D.), soit à partir d'un dictionnaire défini spécifiquement pour les besoins de l'enquête. Il contient au minimum le code désignant la variable, la signification précise de la variable, le libellé précis de chaque modalité dans le cas de variables qualitatives et le code utilisé pour les "données manquantes".

2 - EXPLOITATION DE LA BASE DE DONNEES

Le mode d'extraction des données dépend essentiellement du type de l'outil de requête qui est associé au S.G.B.D. utilisé (langage de requête standardisé comme le SQL, Structured Query Language, et/ou outil de requête interactif et intuitif par menu). Les données extraites sont destinées à être exportées vers les logiciels statistiques.

Encart 5

SEMANTIQUE ET SYMBOLIQUE

Le mot information recouvre deux notions différentes : l'information est un renseignement qui améliore notre connaissance (signifié) ; l'information est une représentation symbolique de ce renseignement (signifiant). La pratique habituelle du langage fait que nous confondons inconsciemment le signifiant et le signifié. Par contre, les machines de traitement de l'information manipulent exclusivement des symboles, d'une manière totalement indépendante de leur signification.

Dans les enquêtes d'épidémiologie animale, l'information est présente à la fois par le signifiant et le signifié sur les questionnaires (dans la mesure où les questions précisent le codage symbolique de l'information).

En revanche, l'enregistrement sur support informatique fait que l'information est véhiculée et traitée par son signifiant (par exemple la modalité '6' pour décrire l'information "vêlage avec embryotomie chez une vache"). Il faut donc être à même de faire et de conserver en permanence la correspondance entre signifiant et signifié pour pouvoir utiliser le renseignement informatisé tout en lui gardant son sens.

Le traitement automatisé de l'information (i.e. l'informatique) nécessite donc des outils syntaxiques et symboliques permettant d'en résoudre les problèmes de représentation

Par exemple, un dictionnaire des variables annexé à la base de données permet de retrouver rapidement le contenu des fichiers, la signification précise de chaque variable donnée, et le libellé de chaque modalité dans le cas de variables qualitatives

Le dictionnaire référence de façon univoque chaque variable ; de ce fait, les risques d'ambiguïté sur les variables analysées sont levés (pas d'erreur possible sur le libellé de la variable, ni sur le sens des modalités), et toutes les éditions nécessaires à l'analyse sont documentées : cela apporte sécurité et confort, compte tenu de la difficulté à désigner le contenu d'une variable avec son seul nom de 8 caractères comme c'est le

cas dans les logiciels statistiques (SAS par exemple). Pour aller plus loin dans l'usage du dictionnaire, on peut l'utiliser en créant automatiquement les ordres de création des fichiers dans le logiciel statistique, ce qui améliore considérablement l'interface entre le S.G.B.D. et le logiciel statistique.

CONCLUSION

La gestion de l'information dans les enquêtes d'épidémiologie animale, pour être menée à bien, doit respecter les règles suivantes : fiabilité des sources qui sont de natures diverses, cohérence des informations produites par celles-ci, bonne connaissance des relations entre les différents types d'information, choix préalable d'un système de gestion de bases de données et d'une stratégie d'analyse des données.

Cette gestion poursuit deux objectifs principaux, obtenir des données standardisées autorisant

l'usage de logiciels statistiques et d'ordinateurs, assurer que les différents processus de transformation qui permettent d'obtenir des données "transformées" à partir des données "collectées" puissent s'effectuer de telle sorte que leur qualité intrinsèque (fiabilité, cohérence) soit conservée.

Cette gestion, organisation du recueil des données et contrôle de leur qualité, demande à être prise en compte d'une manière simultanée et interactive avec le processus de conception/réalisation de l'enquête.

REFERENCES BIBLIOGRAPHIQUES

- Barnouin J. - Enquête écopathologique continue en élevages-observatoires chez les ruminants. Le système de codification et de vérification des données. *Annales de Recherche Vétérinaire*, 1980, 11, 351-366.
- Brunet J. - Epidémiologie-surveillance en abattoir : retour des informations sanitaires en élevage ovin et caprin. *Epidémiologie et Santé Animale*, 1985, 8, 27-41.
- Calavas D. - Coûts des pathologies en élevage bovin laitier : Conception et réalisation de l'enquête - Elaboration du référentiel de données. *"Compte rendu d'étude"*, 10, 1, Centre d'Ecopathologie Animale ed., Villeurbanne, 1992 a, 178 pp.
- Calavas D. - Mammites des brebis allaitantes : Conception de l'enquête. *"Compte rendu d'étude"*, 11, 1, Centre d'Ecopathologie Animale ed., Villeurbanne, 1992 b, 143 pp.
- Calavas D., Rosner G. and Ducrot C. - Collection of data by non professionals in ecopathology survey. *Veterinary Research*, 1994, 25, 98-103.
- Chen P. P. - The entity-relationship model. Toward an unified view of data. *ACM T. Database Systems*, 1976, 1, 9-36.
- Cood E. F. - Further normalization of the database relational model. In RUSTIN R. (Ed), *Courant Computer Science Symposia : Database Systems*, Englewood Cliffs (USA), Prentice-Hall ed., Proceedings, 1972, 33-64.
- Cood E. F. - Recent investigations in relational data base systems. In ROSENFELD J. L. (Ed), *Information Processing 74*, Amsterdam, North Holland, Proceedings, 1974, 1017-1021.
- Curtis C. R. - The marriage of epidemiologic and experimental research : a unified approach. In 7th International Conference on Production Disease in Farm Animals, Ithaca, New York, Proceedings, 1989, 74-87.
- Ducrot C. - Anoestrus des vaches allaitantes. *"Compte-rendu d'étude"*, 12, 1, Centre d'Ecopathologie Animale ed., Villeurbanne, 1992, 98 pp.
- Faye B. et Barnouin J. - Objectivation de la propreté des vaches laitières et des stabulations : l'indice de propreté. *Bulletin Technique CRZV Theix*, 1985, 59, 61-67.
- Faye B., Barnouin J. et Lescourret F. - Objectifs principaux et stratégie de l'enquête écopathologique Bretagne sur la vache laitière. *Epidémiologie et Santé Animale*, 1989, 15, 23-31.
- Faye B., Calavas D. et Rosner G. - La fiabilité des données dans les enquêtes d'écopathologie. *Bulletin de l'Office International des Epizooties*, 1994, 13-3, (accepté pour publication).
- Hall S. A., Dawson P. S. and Davies G. - VIDA II : a computerized diagnostic recording system for veterinary investigation centres in Great Britain. *Veterinary Research*, 1980, 106, 260-264.
- Josse J. - Informatique et écopathologie. *Recueil de Médecine Vétérinaire*, 1986, 162, 819-823.
- Lescourret F., Genest M., Barnouin J., Chassagne M. and Faye B. - Data modeling for database design in production and health monitoring systems for dairy herds. *Journal of Dairy Science*, 1993, 76, 4, 1053-1062.
- Lescourret F., Perochon L., Coulon J. B., Faye B. and Landais E. - Modelling and information system using the MERISE method for agricultural research : the example of a database for a study on performances in dairy cows. *Agricultural Systems*, 1992, 38, 149-173.
- Martin J. - Database organisation. Prentice-Hall ed., Englewood Cliffs, USA, 1977, 185 pp.
- Martin S. W., Meek A. H. and Willeberg P. - *Veterinary Epidemiology - Principles and methods*. Iowa State University Press ed., Ames, Iowa, 1987, 343 pp.
- Monicat F. - Arthrites des caprins. *"Compte rendu d'étude"*, 4, Centre d'Ecopathologie Animale ed., Villeurbanne 1988, 343 pp.
- Peretz G., 1992 - Qualité hygiénique du lait et des fromages de chèvre - Mise en place de l'enquête. *"Compte rendu d'étude"*, 6, 2,

- Centre d'Ecopathologie Animale ed.,
Villeurbanne, 136 pp.
- Philipot J. M. - Vêlage et infécondité des vaches
laitières. "Compte rendu d'étude", 15, Centre
d'Ecopathologie Animale ed., Villeurbanne,
1993, 137 pp.
- Pons C. - Guide pratique SQL. "Presses Pocket",
PSI ed., Paris, 1989, 180 pp.
- Rumeau-Rouquette C., Blondel B., Kaminski M.
et Breart G. - Epidémiologie. Méthodes et
pratique. "Collection Statistiques en biologie et
en médecine", Médecine-Sciences,
Flammarion ed., Paris, 1993, 312 pp.
- Schukken Y. H., Van de Geer D., Grommers F. J.
and Brand A. - Assessing the repeatability of
questionnaire data from dairy farms.
Preventive Veterinary Medicine, 1989, 7, 31-
38.
- Stone D. J. W., Thrusfield M. V. - A small animal
clinical and epidemiological relational
database. Preventive Veterinary Medicine,
1989, 7, 289-302.
- Sulpice P. - Appropriation de l'outil informatique
par des enquêteurs au cours d'une
expérimentation de saisie sur site. In Colloque
Européen Agrimatica "Informatique et
télématique agricoles. Méthodes et conduites
de projets", Charbonnières-les-Bains, Télé
Promotion Rurale Rhône-Alpes ed.,
Proceedings, 1992, 263-271.
- Sulpice P. - Saisie informatique des données en
élevage. "Compte-rendu d'étude", 13, Centre
d'Ecopathologie Animale ed., Villeurbanne,
1993, 87 pp.
- Sulpice P., Bugnard F. and Calavas D. -
Databases save time and improve the quality
of the design, management and processing of
ecopathology surveys. Veterinary Research,
1994, 25, 120-126.
- Tardieu H., Rochfeld A. et Coletti R. - La méthode
MERISE. Tome 1. Principes et outils. Les
Editions d'Organisation ed., Paris, France,
1983, 320 pp.
- Toma B., Bénet J. J., Dufour B., Eloit M., Moutou
F. et Sarana M. - Glossaire d'épidémiologie
animale. Point Vétérinaire ed., Paris, 1991,
365 pp.

