

L'épidémiologie en santé au travail (II)

Étapes d'une étude épidémiologique en milieu professionnel

Cet article, deuxième d'une série de trois, vise à rappeler les grandes étapes d'une étude épidémiologique en milieu professionnel. Le premier article, publié dans le n° 112 des Documents pour le Médecin du Travail (dernier trimestre 2007), rappelait un certain nombre de définitions en épidémiologie [1]. Le troisième article (DMT n° 114, deuxième trimestre 2008) donnera aux lecteurs quelques clés pour la lecture critique d'une publication d'étude épidémiologique analytique en santé au travail.

Cette série d'articles complète l'information du médecin du travail en épidémiologie initiée par le guide « Surveillance épidémiologique de la mortalité et investigation d'agrégats spatio-temporels en entreprise. Principes généraux et données nécessaires », publié par l'INRS et l'InVS en juin 2004.

En résumé

Les grandes étapes d'une étude épidémiologique comprennent :

- la revue de la littérature sur la thématique d'intérêt aboutissant à la formulation d'objectifs et d'hypothèses précis ;
- la définition de la population étudiée et de l'échantillonnage ;
- le calcul du nombre de sujets nécessaire pour répondre à l'objectif fixé ;
- la rédaction du protocole, document de référence utilisé tout au long de l'étude ;
- la prise en compte des aspects éthiques et réglementaires présentés par l'étude ;
- le recueil des données, concernant l'état de santé mais aussi l'exposition professionnelle au facteur de risque étudié ;
- l'analyse statistique ;
- la communication et la publication des résultats, à la communauté scientifique, aux partenaires et aux sujets ayant participé.

Les réponses à ces interrogations ne se situent pas toujours à l'échelle de l'individu mais à celle d'une population de salariés et nécessitent le recours à l'épidémiologie. L'objectif de cet article est de dresser un panorama non exhaustif des différentes étapes d'une étude épidémiologique en milieu professionnel. Ces différentes étapes sont détaillées ci-après.

La revue de la littérature

OBJECTIFS

L'étape bibliographique et la connaissance des enquêtes éventuellement en cours dans les autres pays sont essentielles. Il s'agit de faire le bilan des connaissances le plus complet et le plus récent possible sur la question de manière à :

- ne pas refaire ce qui a déjà été fait (sauf dans le cas où l'objectif est de confirmer des résultats) ;
- se servir des dernières connaissances acquises sur le sujet pour poser une question ou une hypothèse nouvelle sur le sujet, ayant une chance raisonnable d'être testée ;
- obtenir des informations sur la fréquence de la maladie, du facteur d'exposition, sur des facteurs de confusion éventuels (**encadré 1**) ;
- avoir une idée des difficultés rencontrées par d'autres chercheurs et des méthodes utilisables ou non.

E. BOURGKARD,
V. DEMANGE, C. AUBRY

Département Épidémiologie en entreprise,
INRS

Dans leur pratique quotidienne, les médecins du travail sont régulièrement confrontés à des interrogations sur l'état de santé des salariés qu'ils suivent notamment sur les liens entre les pathologies observées et les conditions de travail.

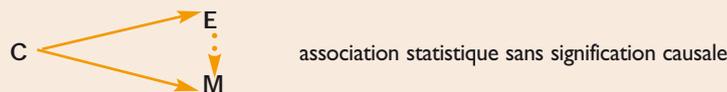
 inrs

Documents pour le Médecin du Travail
N° 113
1^{er} trimestre 2008

Définition et prise en compte d'un facteur de confusion

Un **facteur de confusion** (C) est un tiers facteur qui vient perturber l'étude de la relation entre le facteur d'exposition (E) et la maladie (M) étudiés. Il ne résulte pas de l'intervention de l'investigateur (lors de l'échantillonnage par exemple), mais existe de façon naturelle dans la population. S'il n'est pas pris en compte, il est à l'origine d'un biais de confusion.

C est lié à E et à M et l'on observe une liaison entre E et M qui n'est due qu'à la présence de C.



EXEMPLE HISTORIQUE : Snow a observé, au XIX^e siècle à Londres, un lien entre survenue du choléra et altitude, les quartiers les plus bas situés étant les plus touchés. La qualité de l'air a été d'abord incriminée, étant moins bonne dans les quartiers bas situés. En fait, c'est la qualité de l'eau qui était en cause, les quartiers les plus bas situés recevant également l'eau de moins bonne qualité.



D'après Goldberg M. *L'épidémiologie sans peine*, Ed. Frison-Roche, Paris, 1990

Prise en compte des facteurs de confusion

- à l'échantillonnage
- à l'analyse

Pour pouvoir les prendre en compte il faut les connaître en ayant fait une revue de la littérature.

MOYENS

Il existe des bases de données informatisées des publications (articles, livres) qui sont indexées en fonction de mots-clés. Elles peuvent être retrouvées par des équations de recherche (agencements de mots-clés) bien construites. Il faut également s'intéresser à la « littérature grise » qui correspond selon la Haute autorité de santé à la « production non indexée dans les catalogues officiels d'édition et dans les circuits conventionnels de diffusion de l'information : résumés pour les congrès, documents gouvernementaux, études non publiées, rapports et autres documents non conventionnels » (**encadré 2**).

RÉSULTATS : LE FOND ET LA FORME

Les articles les plus pertinents pour le sujet et ayant une qualité suffisante doivent être sélectionnés (cf. prochain article « *Lecture critique d'une publication d'une étude épidémiologique* »). Leur contenu doit être synthétisé pour répondre aux objectifs de la revue de la littérature. Celle-ci fait partie intégrante du protocole et sera également utilisée pour la rédaction des publications. Elle doit être présentée selon des normes ; la plus utilisée par les revues est la norme dite Vancouver

(ex : AMODEO M., BOUDOT H., DESFRAY F. ET AL. - La coiffure : une enquête de terrain en Côte D'or. *Doc Méd Trav.* 2004 ; 99 : 367-81).

SPÉCIFICITÉS DU MILIEU PROFESSIONNEL : QUELQUES POINTS DE REPÈRE

Un certain nombre de revues françaises ou anglo-saxonnes traitent spécifiquement de la question de la santé au travail :

• françaises

- *Documents pour le Médecin du Travail* :

www.dmt-prevention.fr

- *Hygiène et sécurité du travail* : **www.hst.fr**

- *Travail & Sécurité* : **www.travail-et-securite.fr**

- *Archives des Maladies Professionnelles et de Médecine du Travail* : **www.masson.fr/revues/admp**

- anglophones :

- *Occupational and environmental medicine* :

<http://oem.bmj.com>

- *International archives of occupational and environmental health* : **www.springerlink.com**

- *American journal of industrial medicine* :

www3.interscience.wiley.com/cgi-bin/home

- *Scandinavian journal of work, environment & health* : **www.sjweb.fi**

Il existe également un certain nombre de bases de données qui concernent spécifiquement le milieu professionnel.

• **Expositions :**

- **Fiches toxicologiques :** synthèses des informations disponibles sur les dangers d'une substance ou d'un groupe de substances chimiques (www.inrs.fr/fichetox).
- **DEMETER** (Documents pour l'évaluation médicale des produits toxiques vis-à-vis de la reproduction) : fiches contenant des informations sur les dangers vis-à-vis de la reproduction d'une soixantaine de produits et permettant d'évaluer le risque dû à l'exposition en fonction de sa période de survenue avec des propositions de conduites à tenir.
- **MétroPol :** méthodes de prélèvement et d'analyse de l'air pour l'évaluation de l'exposition professionnelle aux agents chimiques (www.inrs.fr/metropol).
- **Colchic :** données d'expositions professionnelles collectées par l'INRS et les laboratoires inter-régionaux des CRAM.
- **Biotox :** Inventaire des laboratoires effectuant des dosages de toxiques industriels (www.inrs.fr/biotox).
- **EFICATT :** Exposition fortuite à un agent infectieux et conduite à tenir en milieu de travail (www.inrs.fr/eficatt)

• **Maladies professionnelles :** tableaux des maladies professionnelles (www.inrs.fr/mp)

• **Accidents du travail : EPICEA,** base de données nationale et anonyme, non exhaustive, rassemblant plus de 20 000 cas d'accidents du travail survenus, depuis 1990, à des salariés du régime général de la Sécurité sociale (www.inrs.fr/epicea).

• **Réglementations : Perinorm :** base de données bibliographiques de règlements et normes (www.perinorm.com/pol/accueil.php).

• **Générales :**

- base de données bibliographique du National Institut of Occupational Safety and Health (www.cdc.gov/niosh/srchpage.html)
- monographies sur l'évaluation des risques de cancérogénicité pour l'homme du Centre International de Recherche sur le Cancer (www.iarc.fr)
- **INRS biblio** [2] ⁽¹⁾ ;
- Les enquêtes en milieu professionnel (www.travail.gouv.fr/etudes-recherche-statistiques/statistiques/sante-au-travail/87.html) :
 - **Surveillance médicale en milieu professionnel (SUMER) :** Cette enquête, pilotée par la DARES (Direction de l'animation de la recherche, des études et des statistiques) et la

(1) Les différentes procédures de recherche publiées dans la rubrique DMT assistance sont accessibles sur le site www.dmt-prevention.fr

Les bases de données bibliographiques biomédicales

Les bases de données médicales

MEDLINE PASCAL SCISEARCH BIOSIS CINAHL

La Cochrane Library

Cochrane Database of Systematic Reviews (CDSR)
Database of Abstracts of Reviews of Effectiveness (DARE)
Cochrane Controlled Trials Register (CCTR)
Cochrane Review Methodology Database (CRMD)

La littérature grise

Les bases de données de littérature grise

La Banque de données de santé publique (BDSP)
SIGLE (System for Information on Grey Literature in Europe)
HealthSTAR, de la National Library of Medicine (NLM), et l'American Hospital Association
NHS Economic Evaluation Database produite par le NHS Centre for Reviews and Dissemination
HSRProj (Health Services Research Projects in Progress) de la NLM

Les agences et institutions gouvernementales, les sociétés savantes, les associations professionnelles et les agences d'évaluation technologique

Cf. HAS www.has-sante.fr/portail/upload/docs/application/pdf/analiterat.pdf

ENCADRÉ 2



Direction générale du travail (Inspection médicale du travail), décrit les contraintes organisationnelles, les expositions professionnelles de type physique, biologique et chimique auxquelles déclarent être soumis les salariés. Elle a eu lieu en 1987, 1994 et 2003.

- **Enquête décennale de santé 2002-2003** : La mesure de la morbidité et de la consommation de biens médicaux constituent le cœur de l'enquête. Un questionnaire sur les conditions de travail au cours de la vie a été ajouté à cette 4^e enquête décennale de l'Institut national de la statistique et des études économiques (INSEE).

- **Santé et Itinéraire Professionnel (SIP)** : enquête conçue dans le cadre d'un partenariat entre la DREES (Direction de la recherche, des études, de l'évaluation et des statistiques) et la DARES avec l'appui scientifique du Centre d'études de l'emploi. Sa mise en œuvre est assurée par l'INSEE. Elle a pour objectif de mettre en évidence les liens et interactions entre la santé et l'itinéraire professionnel. Pour cela, elle retrace le parcours professionnel et de santé des personnes enquêtées.

- **Enquête décennale de santé 2002-2003** : La mesure de la morbidité et de la consommation de biens médicaux constituent le cœur de l'enquête. Un questionnaire sur les conditions de travail au cours de la vie a été ajouté à cette 4^e enquête décennale de l'Institut national de la statistique et des études économiques (INSEE).

- **Enquête Santé et vie professionnelle après 50 ans (Enquête SVP50)** : Cette enquête, réalisée en 2003 par le groupe « épidémiologie » du Centre inter services de santé et de médecine du travail en entreprise (CISME) auprès de salariés de plus de 50 ans, s'intéresse aux conditions de travail présentes et passées du salarié, à sa perception de la retraite et à divers aspects de son état de santé.

...

La formulation des objectifs et des hypothèses

Lorsque la revue de la littérature met en évidence un manque de connaissances, l'étude doit apporter des éléments de réponse nouveaux. D'une façon générale, une enquête épidémiologique a comme objectif d'apporter une réponse à une question précise. C'est la nature de cette question qui va conditionner le choix de la méthode d'investigation. La formulation de l'objectif doit être extrêmement précise et mentionner quelle est

la population étudiée, la ou les pathologies prise(s) en compte, les facteurs de risques suspectés. Les hypothèses prennent position pour une réponse à la question posée par les objectifs.

En fonction des champs de l'épidémiologie dans lesquels on se trouve, on aura des hypothèses et des objectifs différents.

EXEMPLE DANS UNE ÉTUDE DESCRIPTIVE

« Les objectifs de l'étude présentée ici sont de :

- montrer l'état de santé des salariés en fonction du type de grande surface commerciale et du sexe,
- décrire les conditions de travail et leur retentissement possible sur la santé des salariés,
- rechercher d'éventuelles différences entre supermarchés et hypermarchés,
- mieux comprendre et agir sur des conditions de travail qui paraissent génératrices de pathologies » [3].

Cette étude descriptive peut générer des hypothèses pour rechercher une cause entre différents états de santé (type de grande surface, conditions de travail).

EXEMPLE DANS UNE ÉTUDE ANALYTIQUE

« *The present study was carried out to determine the frequency of ocular and respiratory symptoms in lifeguards working in indoor pools and to examine the extent to which they were related to the measured concentrations of NCl_3 in the pool areas.* » (cette étude a été menée pour déterminer la fréquence des symptômes oculaires et respiratoires chez les maîtres-nageurs travaillant dans des piscines couvertes, et pour évaluer sa relation avec les concentrations en trichlorure d'azote mesurées dans les piscines) [14] : il s'agit bien d'explorer le lien entre une exposition et des effets sur la santé.

EXEMPLE DANS UNE ÉTUDE ÉVALUATIVE

« *A prospective epidemiologic study was conducted in two units (140 beds and 57 nursing assistants) of a nursing home to demonstrate the efficacy of an ergonomic intervention strategy to reduce back stress to nursing personnel.* » (une étude épidémiologique prospective a été menée dans deux unités de soins [140 lits et 57 infirmières] en milieu hospitalier pour démontrer l'efficacité d'une stratégie d'intervention ergonomique pour réduire les rachialgies du personnel soignant [15]). Ici c'est l'hypothèse qui est présentée, c'est-à-dire que les auteurs testent l'effet bénéfique de l'intervention sur les lombalgies.

Population étudiée, échantillonnage, représentativité, puissance et calcul du nombre de sujets nécessaire

On appelle « **population cible** » la population à laquelle on veut généraliser les résultats, et « **population source** » la population d'où est extrait l'échantillon étudié. Si l'enquête concerne l'ensemble de la population, on dit que c'est une **enquête exhaustive**. Le plus souvent, l'enquête est réalisée par sondage sur un échantillon de la population.

L'**échantillonnage** est la façon dont sont extraits les sujets de la population source pour constituer l'échantillon. L'échantillonnage est réalisé le plus souvent par tirage au sort, ce qui garantit que tout sujet appartenant à la population peut faire partie de l'échantillon avec une probabilité connue. Cela permet de garantir la **représentativité** de l'échantillon (échantillon aléatoire) pour certaines caractéristiques (la répartition par sexe, par âge, par catégorie socio-professionnelle...) par rapport à la population dont il est extrait. Cela permettra d'extrapoler les résultats obtenus sur l'échantillon à la population (si toute la suite de l'étude réalisée est de qualité).

Cet échantillonnage peut être simple ou élémentaire : l'échantillon est tiré au sort dans la base de sondage en une seule opération. Le sondage peut aussi être **stratifié** : la population de base est divisée en sous-groupes homogènes pour une caractéristique et l'échantillonnage est effectué au sein de chacun de ces sous-groupes ou strates. Enfin le sondage est dit **en grappes** si le tirage au sort est effectué non pas sur des sujets, mais sur des sous-groupes (grappes) : dans ce cas, tous les individus du sous-groupe tiré au sort participent à l'enquête.

Exemple : on souhaite réaliser un échantillon d'enseignants de collègues pour une étude épidémiologique descriptive. Compte-tenu du nombre élevé de professionnels concernés, on décide de réaliser un échantillonnage. Voici trois modalités possibles :

- tirage au sort parmi la liste de tous les enseignants de collègues : enquête par sondage élémentaire ;
- tirage au sort d'un échantillon d'enseignants dans chaque discipline : stratification sur la discipline ;
- tirage au sort de collègues et inclusion de tous les enseignants des collègues sélectionnés : sondage en grappes.

Dans la réalité, on couple souvent ces méthodes pour réaliser l'échantillonnage.

La composition de l'échantillon peut différer de celle de la population du fait du hasard, cela s'appelle les fluc-

tuations d'échantillonnage. Il faut les prendre en compte à l'aide de méthodes statistiques comme l'estimation. L'estimation permet d'évaluer à partir de la valeur observée dans l'échantillon la véritable valeur de cette variable dans la population. Du fait de l'incertitude liée aux fluctuations d'échantillonnage, on ne peut que calculer une probabilité qu'à la vraie valeur de se trouver dans un intervalle. On dira par exemple : la fréquence de la maladie M dans la population est de 15 %, intervalle de confiance à 95 % : 13 %-17 % ou le nombre de paquets-années chez les fumeurs est en moyenne de 10,5 avec un intervalle de confiance à 95 % : 8-13. Cela veut dire que l'on a 5 % de risque de faire une évaluation erronée, c'est-à-dire que la fréquence ou la moyenne se trouvent en dehors de l'intervalle de confiance.

Calcul du nombre de sujets nécessaire : il s'agit de l'effectif minimum suffisant de l'échantillon pour obtenir une précision donnée avec un risque d'erreur donné. Il dépend des ces 2 paramètres que l'on choisit ainsi que de la variabilité du paramètre étudié dans la population (plus sa variabilité est grande, plus il faudra de sujets pour obtenir une estimation précise). Cette variabilité peut soit être retrouvée par la revue de la littérature, soit devra être estimée par une étude pilote.

Puissance de l'étude : l'estimation peut être très imprécise du fait d'un manque de puissance de l'étude qui dépend notamment du nombre de sujets de l'échantillon. Plus on aura de sujets dans l'échantillon, plus la puissance de l'étude et plus la précision de l'estimation seront grandes. Si l'on incluait tous les sujets dans l'échantillon, la précision serait optimale (aux erreurs ou approximations de mesure près) et l'intervalle de confiance nul. Mais le gain de précision devient négligeable au delà d'un certain effectif, compte-tenu du coût du recueil des informations.

La rédaction du protocole

Obligatoire, le protocole écrit sert de base scientifique (état des connaissances sur le sujet, objectifs et hypothèses de l'étude), technique (identification des instruments de mesure, façon de recueillir les données), réglementaire (autorisations demandées), contractuelle (relations entre promoteur et investigateurs, répartition des tâches, règles de communication des résultats) puis de fil conducteur pendant toute la durée de l'étude. Il est fondamental s'il existe plusieurs investigateurs mais également s'il n'y en a qu'un seul (il évite les dérives dans l'espace et dans le temps). Le protocole doit envisager tous les cas de figure concrets pouvant se présenter pendant le déroulement de l'enquête et la façon précise de les traiter, notamment les

modalités de prise en compte des non répondants et des perdus de vue.

La participation à une étude repose sur le volontariat du sujet. Si certains sujets « éligibles », c'est-à-dire répondant aux critères d'inclusion définis, refusent de participer, il est important de savoir s'ils diffèrent des sujets qui acceptent de participer. Le nombre de non répondants est limité au maximum. Si leur nombre est important, une enquête complémentaire pourra être menée en tirant au sort des non répondants pour connaître les informations manquantes.

Les perdus de vue sont des sujets qui ont commencé à participer à une enquête exposés-non exposés, impliquant un suivi au cours du temps et qui ne participent plus à l'étude sans que l'on sache pourquoi. Là aussi, une enquête complémentaire peut être réalisée en tirant au sort des sujets perdus de vue afin d'obtenir des informations manquantes.

Si la raison du refus de participer à l'étude ou de ne plus participer à l'étude est liée à la maladie étudiée ou au facteur d'exposition, les résultats obtenus sur l'échantillon ne seront pas valides pour la population cible (on parle d'un biais, cf. prochain article à paraître dans le *DMT* n° 114) et on ne pourra pas généraliser les résultats obtenus sur l'échantillon à cette population. Classiquement, les caractéristiques dont on dis-

pose entre ces 2 groupes sont comparées. S'il existe une différence statistiquement significative, les résultats ne pourront être généralisés à la population cible.

Spécificités de l'épidémiologie en milieu professionnel : le protocole est un contrat entre les investigateurs, promoteurs et le dirigeant de l'entreprise. L'**encadré 3** présente un exemple de plan de protocole, à adapter au cas particulier de chaque étude.

Les aspects éthiques et réglementaires

BREF HISTORIQUE

La prise de conscience de la nécessité de la protection des personnes apparaît au procès de Nuremberg avec la création du Code de Nuremberg. Une éthique des médecins et chercheurs doit être développée. Quelques années plus tard, la déclaration d'Helsinki de 1964 fait suite au constat que les principes éthiques ne sont pas respectés. En France, la législation évolue de

ENCADRÉ 3

Exemple de plan de protocole utilisé au département Épidémiologie en entreprise de l'INRS, inspiré de l'ouvrage « *Epidémiologie - Principes et méthodes quantitatives* » de BOUYER et al [16].

1. Titre
2. Résumé
3. Contexte de l'étude
4. Organisation générale : coordonnées des organisateurs et participants à l'étude
5. Objectifs de l'étude
6. Méthodes
 - Type d'étude
 - Faisabilité et puissance statistique
 - Définition des groupes étudiés
 - Données recueillies
 - Mode de recueil et gestion des données
 - Analyse statistique
7. Chronologie prévisionnelle
8. Mesures réglementaires et éthiques
9. Communications et publications
10. Références bibliographiques
11. Annexes (à adapter)
 - Dates rédaction et approbation
 - Signatures des organisateurs et participants
 - Avenant(s) signé(s) par les signataires du protocole
 - Lettre d'information du sujet
 - Formulaire de consentement

L'ADEL (Association des épidémiologistes de langue française) présente un plan de protocole plus détaillé dans ses recommandations « *Déontologie et bonnes pratiques en épidémiologie* » [17].

Définitions (article L1121-1 du Code de la Santé publique, Loi du 9 août 2004)

Recherches biomédicales : recherches organisées et pratiquées sur l'être humain en vue du développement des connaissances biologiques et médicales.

Promoteur : personne physique ou morale qui prend l'initiative de la recherche biomédicale, en assure la gestion et vérifie que son financement est prévu. Il est le garant de l'initiative de la recherche et de ses conséquences avec obligation d'assurance.

Investigateur(s) : personne(s) physique(s) qui dirige(nt) et surveille(nt) la réalisation de la recherche. C'est un « médecin justifiant d'une expérience appropriée ».

Autorisation de lieu : les recherches biomédicales ne peuvent être réalisées que dans un lieu disposant des moyens humains, matériels et techniques adaptés à la recherche et compatibles avec les impératifs de sécurité des personnes qui s'y prêtent. Ce lieu doit être autorisé à cet effet pour une durée déterminée.

Principes de protection des personnes (article L1121-2, Loi du 9 août 2004)

Aucune recherche biomédicale ne peut être effectuée sur l'être humain :

- si elle ne se fonde pas sur le dernier état des connaissances scientifiques et sur une expérimentation préclinique suffisante ;
- si le risque prévisible encouru par les personnes qui se prêtent à la recherche est hors de proportion avec le bénéfice escompté pour ces personnes ou l'intérêt de cette recherche ;
- si elle ne vise pas à étendre la connaissance scientifique de l'être humain et les moyens susceptibles d'améliorer sa condition ;
- si la recherche biomédicale n'a pas été conçue de telle façon que soient réduits au minimum la douleur, les désagréments, la peur et tout autre inconvénient prévisible lié à la maladie ou à la recherche, en tenant compte particulièrement du degré de maturité pour les mineurs et de la capacité de compréhension pour les majeurs hors d'état d'exprimer leur consentement. L'intérêt des personnes qui se prêtent à une recherche biomédicale prime toujours les seuls intérêts de la science et de la société.

La recherche biomédicale ne peut débuter que si l'ensemble de ces conditions sont remplies. Leur respect doit être constamment maintenu.

façon significative avec la loi du 20 décembre 1988 relative à la protection des personnes qui se prêtent à des recherches biomédicales, dite « loi Huriot-Sérusclat ». La Directive européenne du 4 janvier 2001 concernant les « essais cliniques de médicaments à usage humain » va être à l'origine de nouvelles évolutions de l'encadrement des recherches biomédicales et sera transposée en droit français le 9 août 2004 dans la « loi de santé publique ». Des décrets d'application (en date du 27 août 2006) ont été publiés au *Journal Officiel*.

CONCEPTS DE BASE SUR LA PROTECTION DES PERSONNES QUI SE PRÊTENT À DES RECHERCHES BIOMÉDICALES

(encadré 4)

Les anciens Comités consultatifs de protection des personnes qui se prêtent à des recherches biomédicales (CCPPRB) sont devenus les Comités de protection des personnes (CPP) dont la composition a été élargie et les compétences étendues. Le régime de déclaration antérieur est remplacé par un régime d'autorisation : toute recherche biomédicale nécessite désormais, pour être mise en œuvre, à la fois un avis favorable d'un CPP et une autorisation de l'autorité compétente qui est, selon le domaine concerné, la Direction générale de la santé (DGS) ou l'Agence française de sécurité sanitaire des produits de santé (AFSSAPS). Leurs demandes et leurs recommandations doivent être suivies d'effet, avec suivi tout au long des protocoles. Tous les protocoles sont évalués sur la balance bénéfique/risque c'est-à-dire sur « le caractère satisfaisant de l'évaluation des bénéfices et des risques attendus ». Le CPP émet un avis qui vaut autorisation.

Les recherches biomédicales ne peuvent être effectuées que sous la direction et la surveillance d'un médecin justifiant d'une expérience appropriée. Un sujet ne peut participer à une étude que s'il a fourni un consentement libre, éclairé et exprès (signature d'un formulaire de consentement après information en s'assurant que celle-ci a été comprise).

DÉCLARATION À LA COMMISSION NATIONALE DE L'INFORMATIQUE ET DES LIBERTÉS (CNIL)

Protection des personnes

Dans le cadre des études épidémiologiques, beaucoup d'informations sont traitées : données d'état civil, causes de décès, données cliniques et biologiques, pathologies... La collecte et le traitement de ces données sont réglementés en France (loi du 6 janvier

1978 relative à l'informatique, aux fichiers et aux libertés modifiée par la loi n° 2004-801 du 6 août 2004). Ainsi, les projets de recherche nécessitant d'une part, le recueil et la transmission à l'organisme de recherche de données directement ou indirectement nominatives et d'autre part, le recours à des moyens organisés pour le traitement de données sont concernés par les procédures de déclaration de traitements de données définies dans cette loi.

La loi du 1^{er} juillet 1994 institue un régime de contrôle spécifique pour les traitements informatisés comprenant des données médicales ayant pour fin la recherche dans le domaine de la santé. La procédure de contrôle comprend deux étapes :

- recueil de l'avis du Comité consultatif sur le traitement de l'information en matière de recherche dans le domaine de la santé. Ce comité se prononce sur « *la méthodologie de la recherche au regard des dispositions de la présente loi, la nécessité du recours à des données nominatives et la pertinence de celles-ci par rapport à l'objectif de la recherche* » (article 40.2) ;

- recueil de l'autorisation de la Commission nationale de l'informatique et des libertés (CNIL) basée sur le dossier adressé au Comité et sur des informations relatives à l'information des personnes et aux aspects informatiques.

Par ailleurs, dans le cadre de la protection des individus, la réglementation prévoit :

- information des personnes participant à l'étude sur la nature des données transmises, les objectifs de l'étude et l'utilisation prévue des données, les destinataires de ces données et leur droit d'accès et de rectification des données les concernant, le droit d'opposition ;
- signature d'un formulaire de consentement éclairé.

Protection des données nominatives

Un projet de recherche doit être réalisé dans des conditions de garantie de protection des données nominatives. Toute donnée ou information collectée en cours d'étude doit être rendue anonyme le plus rapidement possible sauf dans la mesure où l'identification des personnes est indispensable à la conduite de l'étude.

Des procédures spécifiques doivent être mises en place pour assurer l'impossibilité d'identifier, directement ou indirectement, l'individu auquel les données se rapportent, ainsi que la sécurité des informations et de leur traitement, qu'elles soient informatisées ou non.

Des demandes d'autorisation sont à faire avant de commencer toute étude biomédicale :

- au représentant de l'État dans la région pour l'autorisation de lieu ;

- au CPP : avis rendu dans un délai de 35 jours après la date de réception de la demande ;

- à l'autorité compétente de l'État (Direction générale de la santé) : délai de 60 jours après la réception de la demande (au-delà l'absence de réponse vaut autorisation) ;

- au Comité consultatif sur le traitement de l'information en matière de recherche dans le domaine de la santé (CCTIRS) (délai de 30 jours après la réception de la demande, au-delà l'absence de réponse vaut avis favorable) ;

- à la CNIL : (joindre l'avis du CCTIRS : délai de 60 jours après la réception de la demande, au-delà l'absence de réponse signifie un refus).

Le recueil des données en milieu professionnel

(encadré 5)

Les données recueillies ne doivent concerner que ce qui a été défini dans le protocole.

- **Mesure de l'exposition** : différentes méthodes existent et le choix d'une méthode dépend de la nature de l'étude réalisée, de son objectif, des conditions de réalisation :

- **cursus laboris** : il s'agit de reconstituer toute la carrière professionnelle du sujet le plus précisément possible, avec la date de début et de fin de chaque période d'emploi, le type de tâches exécutées ;

- **matrice emplois-expositions** : il s'agit d'un tableau avec en colonnes des intitulés d'emploi et en ligne les nuisances considérées. Dans chaque cellule se trouvent des indices d'exposition (de type exposé/non exposé, probabilité d'exposition, indices de niveau d'exposition, période). Des matrices emplois-expositions ont été spécifiquement développées pour des études épidémiologiques, comme par exemple une matrice emplois-expositions dans l'industrie productrice des métaux durs en France [18] pour étudier le risque de cancer du poumon des travailleurs de cette industrie exposés au carbure de cobalt et au tungstène [19]. Ces matrices conçues spécifiquement pour des études (exposition ciblée, liste de professions non exhaustive) peuvent être difficilement utilisées pour d'autres études. C'est pourquoi l'Institut de veille sanitaire (InVS) a mis en place le programme Matgéné (matrices en population générale) visant à élaborer plusieurs matrices multi-nuisances, adaptées à la population française, dans l'objectif de fournir des données d'exposition professionnelle dans le

Élaboration d'un questionnaire

Un outil spécifique de recueil des données devra être élaboré, s'il n'existe pas déjà de questionnaire validé recueillant les données d'intérêt.

Type de questionnaire : auto-administré (ou auto-questionnaire : le sujet le remplit seul), administré en face-à-face, entretien plus ou moins dirigé (le sujet s'exprime et est plus ou moins guidé par l'enquêteur).

Type de questions : ouvertes (ex : « Dans quels secteurs d'activités avez-vous travaillé ? ») ou fermées (ex : « Dans quels secteurs d'activités avez-vous travaillé ? 1. santé 2. industrie 3. commerce »). Les questions ouvertes, plus difficiles à analyser, sont utilisées plutôt dans une phase exploratoire (à la recherche d'idées nouvelles). Pour les questions fermées, il ne faut pas oublier de modalités ; le plus souvent on conclut la question par une modalité « autre » que l'on peut faire préciser en clair.

Certaines informations subjectives peuvent être évaluées à l'aide d'échelles : la plus connue est l'échelle visuelle analogique pour la douleur. Le sujet positionne un curseur sur une réglette de zéro (aucune douleur) à 10 cm (douleur extrême). L'intensité de la douleur correspond à la distance mesurée sur la réglette.

Les difficultés à connaître

- Le langage doit être compréhensible par n'importe qui. Éviter le langage technique, les abréviations, d'autant plus s'il s'agit d'un auto-questionnaire,
- éviter les doubles questions (« Avez-vous travaillé dans l'industrie et dans le commerce ? »), les doubles négations (« N'avez-vous jamais travaillé dans aucune des entreprises suivantes ? »),
- éviter les questions influençant la réponse (« Pensez-vous qu'il faille interdire le travail de nuit ? »),
- dans les questions fermées, si l'on propose un nombre pair de modalités, cela oblige les personnes à prendre position.

Par exemple : « Êtes-vous satisfait de vos horaires de travail ? »

Nombre pair de modalités :

1. satisfait 2. plutôt satisfait 3. plutôt pas satisfait 4. pas satisfait

Nombre impair de modalités :

1. satisfait 2. moyennement satisfait 3. pas satisfait

Ici, les personnes pourront ne pas prendre position en répondant 2. moyennement satisfait

Tester impérativement le questionnaire auprès de personnes ne connaissant pas le sujet, le modifier ensuite si besoin.

domaine de la surveillance ou de la recherche en santé au travail [10]. Par exemple dans la matrice poussières de farine, élaborée par l'InVS, on peut lire que les artisans boulangers, pâtisseries (entreprises de 3 à 9 salariés) ont une probabilité d'exposition d'indice 4 (plus de 90 % de personnes effectivement exposées aux poussières de farine dans cet emploi), un niveau d'exposition élevé (entre 5 et 10 mg/m³) et connaissent des pics d'exposition, alors que les charcutiers ont une probabilité d'exposition d'indice 2 (entre 10 et 50 % de personnes effectivement exposées aux poussières de farine dans cet emploi), un niveau d'exposition moyen entre 2,5 et 5 mg/m³ et ne subissent pas de pic d'exposition. Quand ces matrices sont appliquées avec des *curvus laboris* individuels, les expositions sont attribuées automatiquement aux individus en fonction de leurs intitulés d'emploi. Les matrices ne permettent pas de tenir compte de la variabilité des expositions à l'intérieur d'un emploi et peuvent

entraîner des erreurs de classement ;

- **métrologie d'ambiance** : on peut choisir de mesurer le niveau d'exposition sur un poste de travail fixe, comme par exemple la mesure de la concentration qui correspond à la fraction conventionnelle inhalable de poussières de bois dans une menuiserie au poste « scie ». Cela est plus précis mais ne correspond pas à l'exposition réelle du sujet, qui peut ne venir que ponctuellement au poste « scie » ou qui n'inhalera pas la totalité des poussières mesurées (particules trop grosses) ;

- **métrologie individuelle** : il s'agit de mesurer la concentration correspondant à la fraction conventionnelle inhalable de poussières de bois par le sujet en plaçant un filtre dans sa zone respiratoire, cela permet de mesurer l'exposition réelle du sujet même s'il change de postes dans la menuiserie ;

- **marqueurs biologiques de l'exposition** : ils sont le reflet de l'absorption de la substance par l'or-

ganisme et permettent de prendre en compte d'autres voies d'absorption que l'inhalation, en complément des mesures atmosphériques. Par exemple, l'exposition au benzène des mécaniciens a été évaluée par des mesures de concentrations atmosphériques du benzène et urinaires du métabolite du benzène, l'acide muconique, en fin de journée de travail, pour avoir une idée de la voie d'absorption cutanée [11].

- **Mesure de l'état de santé**, à adapter selon le type d'étude réalisée, l'effet sur la santé étudié, les conditions de réalisation :

- **interrogatoire** : recherche d'une certaine chronologie des symptômes : amélioration ou disparition pendant les week-end ou les congés, réapparition à la reprise du travail ou après celle-ci,
- **examen clinique** adapté aux symptômes ou pathologies étudiés,
- **examens paracliniques et tests fonctionnels** adaptés aux symptômes ou pathologies étudiés.

- **Recueil d'autres données** d'intérêt pour l'étude définies notamment lors de la revue de la littérature : facteurs de confusion comme l'âge, le sexe, les antécédents médicaux, la catégorie socio-professionnelle...

L'analyse statistique

La validation des données est une étape fondamentale. Il faut rechercher d'éventuelles erreurs (de remplissage des manuscrits, de saisie) ou des données manquantes pour éventuellement les compléter.

Il existe 2 grands types de variables : les variables **quantitatives** (ex : concentration de poussières totales dans l'atmosphère en mg/m^3) et les variables **qualitatives** (ex : salarié en CDI : oui/non).

Parmi les variables quantitatives, on distingue les variables continues qui peuvent prendre un nombre infini de valeurs, qui résultent d'une mesure (ex : concentration de poussières totales dans l'atmosphère en mg/m^3 : $50 mg/m^3$) et les variables quantitatives discrètes qui ne peuvent prendre qu'un nombre fini de valeurs et résultent d'un dénombrement (ex : nombre d'emplois occupés dans une carrière professionnelle : 5).

Parmi les variables qualitatives, on distingue les variables binaires qui n'ont que 2 modalités (ex : salarié en intérim : oui/non), les variables nominales qui peuvent prendre plus de 2 modalités (ex : catégorie socio-professionnelle) et les variables ordinales avec la notion d'un ordre entre les modalités (ex : exposition à un solvant : très forte/ forte /moyennement forte/faible). Une variable quantitative peut toujours être transformée en variable qualitative, mais l'inverse est faux. On perd alors de l'in-

Fig. 1 : Exemple de diagramme en bâtons : Localisation des symptômes cutanés chez des coiffeurs en Côte-d'Or.

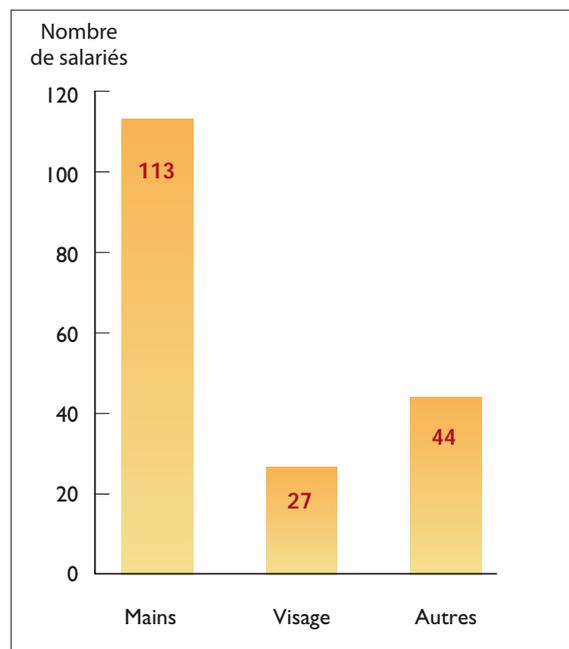


Figure extraite de Amodeo M [12]. La représentation graphique permet une compréhension immédiate par rapport à un tableau. Ici, on voit immédiatement que la localisation prédominante des symptômes cutanés se trouve au niveau des mains.

formation.

La première étape de l'analyse statistique correspond à la **description de l'échantillon** étudié. Elle pourra comporter des représentations graphiques et des tableaux (**figures 1 à 4**) : ils doivent être numérotés, avoir un titre et ne pas nécessiter de recourir au texte pour les comprendre. Les variables qualitatives peuvent être représentées par des diagrammes en bâtons, des diagrammes sectoriels (camembert). Les variables quantitatives continues peuvent être représentées par des histogrammes (dans un histogramme, la surface des rectangles est proportionnelle à l'effectif et les rectangles sont adjacents pour bien faire apparaître le caractère continu de la variable) et, quand elles sont discrétisées, par des diagrammes en bâtons (la hauteur est proportionnelle à l'effectif) ou des camemberts (l'angle au centre est proportionnel à l'effectif).

Il existe 2 types de paramètres statistiques : les **paramètres de position** (moyenne, médiane, quartiles, déciles, percentiles) qui sont des valeurs centrales qui résument l'ensemble des valeurs et les **paramètres de dispersion** (écart-type, écart interquartile) qui caractérisent la variabilité des valeurs unitaires autour de la valeur centrale.

Les valeurs d'une variable quantitative peuvent se distribuer selon certaines règles appelées loi de distri-

Fig. 2 : Exemple de diagramme en bâtons plus compliqué : variation du VEMS* (en ml/an) par rapport aux sujets non exposés en fonction des types d'exposition et des niveaux d'exposition (faible ou forte) chez les sujets ex-fumeurs (étude chez des sujets de la cohorte européenne ECRHS [European Community Respiratory Health Survey], ayant pour objectif la recherche de relations entre expositions professionnelles et fonction respiratoire).

* VEMS : le volume expiratoire maximal seconde est le volume maximal soufflé lors de la première seconde de l'expiration forcée.

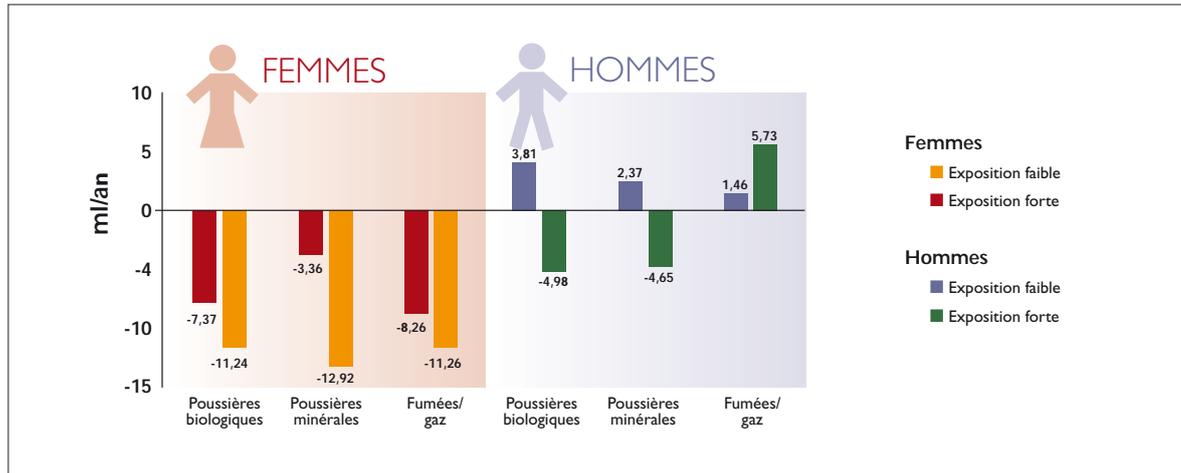


Figure extraite de Matrat M [13], mettant en évidence que, pour les ex-fumeurs, la perte annuelle de VEMS pour les femmes exposées, quel que soit le niveau de l'exposition, est plus importante que pour les hommes par rapport à des non exposés ex-fumeurs de même sexe.

Fig. 3 : Exemple de diagramme sectoriel : Répartition des types de contrat chez des coiffeurs en Côte-d'Or.

Fig. 4 : Association prédite entre l'exposition aux endotoxines après log-transformation et le taux annuel de baisse du VEMS (avec l'écart-type), ajusté sur l'âge, le VEMS de base et les paquets-années.

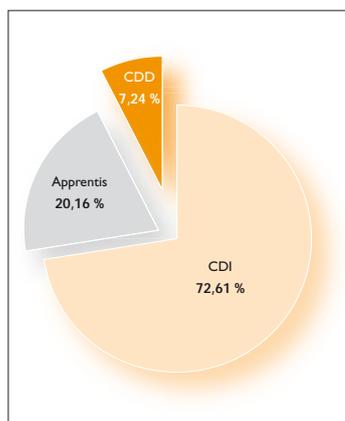


Figure extraite de Amodeo M [12], montrant que le contrat à durée indéterminée (CDI) est majoritaire chez les coiffeurs de l'étude, suivi du contrat d'apprentissage puis du contrat à durée déterminée (CDD).

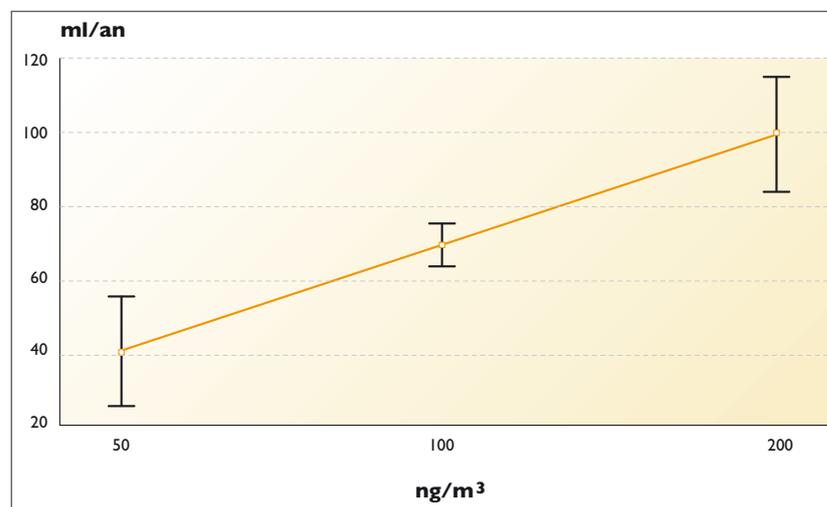


Figure extraite de Matrat M [14]. Cette étude a suivi la fonction pulmonaire d'éleveurs de porcs pendant 3 ans et a étudié la relation avec l'exposition aux endotoxines [14]. Sur la figure, on peut voir que plus l'exposition augmente, plus la baisse annuelle du VEMS augmente après ajustement sur l'âge, le VEMS de base, et les paquets-années.

bution. La plus connue est la loi normale (ou loi de Laplace-Gauss). L'équation de la courbe de fréquence d'une distribution normale ne dépend que de la moyenne et de l'écart-type. Quand une variable dépend de nombreuses autres variables, sa distribution est souvent normale. D'autre part, la loi normale a des propriétés intéressantes c'est pourquoi on essaie toujours de se ramener à une loi normale lorsque l'on fait des sta-

tistiques (par des transformations mathématiques).

L'analyse univariée consiste à analyser séparément les données selon plusieurs variables, dont la variable d'intérêt principal (le facteur d'exposition) pour rechercher d'éventuels facteurs de confusion ou d'interaction. C'est un des moyens utilisés pour sélectionner les variables qui permettront de construire le modèle multivarié.

L'analyse multivariée ou modélisation permet

d'expliquer la survenue de la maladie par l'agencement mathématique (ou modélisation) de différents facteurs dont la variable d'intérêt principal (le facteur d'exposition) et les autres variables explicatives (dont les facteurs de confusion). Il s'agit de tenter de reproduire les données de santé recueillies par une équation mathématique incluant les autres variables recueillies. Cela permet de savoir si le facteur d'exposition augmente le risque de survenue de la pathologie, tout autre facteur étant égal par ailleurs. La quantification de l'association entre facteur d'exposition et maladie se fait par le calcul d'un odds ratio (enquête cas-témoin) ou d'un risque relatif (enquête de cohorte). On parle d'odds ratio ou risque relatif ajusté quand sont prises en compte d'autres variables dans le modèle. L'odds ratio et le risque relatif s'interprètent avec leur intervalle de confiance à 95%.

Exemple : *une étude a suivi des femmes employées dans le secteur du nettoyage et dans le secteur administratif pendant plusieurs années. Le risque de développer un asthme était significativement plus important chez les premières (risque relatif ajusté sur l'âge = 1,5, intervalle de confiance à 95% [1,43-1,57]) [15].*

Une relation dose-réponse peut également être recherchée, à savoir que plus l'exposition augmente, plus l'effet sur la santé est grand. La mise en évidence d'un tel effet (s'il ne résulte pas de biais) est un argument en faveur d'un lien de causalité entre le facteur d'exposition et l'effet sur la santé.

La communication et la publication des résultats

La communication des résultats fait partie des bonnes pratiques en épidémiologie [7].

Les résultats de l'étude doivent être restitués globalement et sous une forme compréhensible aux personnes qui ont participé. L'organisation de cette communication est à prévoir dès le début de l'étude, de manière à pouvoir en informer les sujets au moment de leur inclusion.

D'autre part, une valorisation des résultats sous forme de publication scientifique, internationale ou pas selon le type des travaux, sera prévue, et il est recommandé d'établir dans le protocole qui publiera, l'ordre des noms d'auteurs, sur quelle partie du sujet et dans quelle revue. De même des communications orales dans des congrès scientifiques ou des manifestations professionnelles seront envisagées.

Les résultats, même « négatifs » (absence de mise en évidence de lien entre facteur d'exposition et effet sur la santé) sont dignes d'intérêt et méritent d'être publiés [7]. Ils participent également à la connaissance du sujet et peuvent faire progresser la recherche, dans le sens où les résultats d'une seule étude ne peuvent permettre de conclure la recherche mais alimentent le débat scientifique.

Points à retenir

- La revue de la littérature est fondamentale pour poser un objectif de recherche précis et pertinent. Il existe des bases de données spécifiques au milieu professionnel.
- La formulation des objectifs dépend de la branche de l'épidémiologie concernée par l'étude.
- L'échantillonnage est la façon dont sont extraits les sujets de la population pour constituer l'échantillon.
- Le protocole est le document écrit obligatoire scientifique, technique, réglementaire et contractuel qui fait référence pendant toute la durée de l'étude.
- La recherche biomédicale est réglementée sur le plan de la protection des personnes et des données nominatives. Le délai d'obtention des autorisations réglementaires doit être prévu par le calendrier de l'étude.
- La mesure de l'exposition peut se faire de différentes façons, dépendant de l'objectif de l'étude et aussi des conditions pratiques de réalisation de l'étude : cursus laboris, matrice emplois-expositions, métrologie d'ambiance ou individuelle, marqueurs biologiques de l'exposition.
- Les résultats de l'étude doivent être restitués globalement et sous une forme compréhensible aux personnes qui ont participé. Même en l'absence de mise en évidence de lien entre facteur d'exposition et effet sur la santé, les résultats doivent être valorisés et mis à la disposition de la communauté scientifique.

Bibliographie

- [1] BOURGKARD E., DEMANGE V., AUBRY C - L'épidémiologie en santé au travail (I) : Définitions et concepts. *Doc Méd Trav*, 2007 ; 112 : 477-86.
- [2] BJAOUI A., DEVEL S. - La base de données INRS Biblio. Prévention des risques professionnels. *Doc Méd Trav*, 2007 ; 111 : 411-16.
- [3] DOUSSON C, FERRAND C, GROSSETETE A, BIERME J ET AL. - Etat de santé des salariés de la grande distribution : Epigrandis, une étude descriptive dans le département du Rhône. Etudes et enquêtes TF 111. *Doc Méd Trav*. 2002 ; 89, 1^{er} trimestre 2002 : 29-49.
- [4] MASSIN N, BOHADANA AB, WILD P, HÉRY M ET AL. - Respiratory symptoms and bronchial responsiveness in lifeguards exposed to nitrogen trichloride in indoor swimming pools. *Occup Environ Med*. 1998 ; 55 (4) : 258-63.
- [5] GARG A, OWEN B - Reducing back stress to nursing personnel: an ergonomic intervention in a nursing-home. *Ergonomics*. 1992 ; 35 (11) : 1353-75.
- [6] BOUYER J, HÉMON D, CORDIER S, DERRIENIC F ET AL. - Epidémiologie. Principes et méthodes quantitatives. Paris : les éditions INSERM ; 1993 : 498 p.
- [7] Déontologie et bonnes pratiques en épidémiologie. Révision des recommandations de 1998. Proposition du groupe de travail ad hoc de l'ADELF. Bordeaux : Association des Epidémiologistes de Langue Française ; 2003 : 32 p.
- [8] MOULIN JJ, ROMAZINI S, LASFARGUES G, PELTIER A ET AL. - Elaboration d'une matrice emplois-expositions dans l'industrie productrice de métaux durs en France. *Rev Epidemiol Santé Publique*. 1997 ; 45 (1) : 41-51.
- [9] MOULIN JJ, PERDRIX A, LASFARGUES G, ROMAZINI S ET AL. - Etude épidémiologique de mortalité dans l'industrie productrice de métaux durs en France. Note documentaire ND 2057. *Cah Notes Doc*. 1997 ; 168, 3^e trimestre 1997 : 411-28.
- [10] LUCE D, FÉVOTTE J, ARSLAN M, DELABRE L ET AL. - Le programme Matgéné. Matrices emplois-expositions en population générale. État d'avancement - septembre 2005. Saint-Maurice : InVS ; 2006 : 60 p.
- [11] GAUDIN R, DUCOS P, FRANCIN JM, MARSAN P ET AL. - Exposition au benzène chez les mécaniciens. Evaluation atmosphérique et surveillance biologique. Note documentaire ND 2174. *Cah Notes Doc*. 2002 ; 188, 3^e trimestre 2002 : 27-36.
- [12] AMODEO M, BOUDOT H, DESFRAY F, DUCROT-HENRY L ET AL. - La coiffure : une enquête de terrain en Côte-d'Or: Etudes et enquêtes TF 133. *Doc Méd Trav*. 2004 ; 99, 3^e trimestre 2004 : 367-81.
- [13] MATRAT M, ANDUJAR P, PAIRON JC - Maladies respiratoires professionnelles. Etudes et enquêtes TF 156. *Doc Méd Trav*. 2006 ; 108, 4^e trimestre 2006 : 495-505.
- [14] VOGELZANG P, VAN DER GULDEN JW, FOLGERING H, KOLK JJ ET AL - Endotoxin exposure as a major determinant of lung function decline in pig farmers. *Am J Respir Crit Care Med*. 1998 ; 157 (1) : 15-18.
- [15] KARJALAINEN A, MARTIKAINEN R, KARJALAINEN J, KLAUKKA T ET AL. - Excess incidence of asthma among Finnish cleaners employed in different industries. *Eur Respir J*. 2002 ; 19 (1) : 90-95.