

Organisation de collecte et d'appariement de données pour la surveillance épidémiologique des travailleurs potentiellement exposés aux nanomatériaux en France (EpiNano)

Delphine Jezewski-Serra¹, Laurène Delabre¹, Stéphane Ducamp^{1, 2}, Yuriko Iwatsubo¹, Irina Guseva-Canu¹

Résumé

Le dispositif EpiNano vise à surveiller les travailleurs potentiellement exposés aux nanomatériaux intentionnellement produits en France. Ce dispositif est basé, d'une part, sur les données d'hygiène industrielle recueillies au sein des entreprises afin de caractériser l'exposition aux nanomatériaux des postes de travail et d'autre part celles issues d'auto-questionnaires complétés par les participants. Ces données seront ensuite appariées aux données de santé des bases médico-administratives nationales (suivi passif des événements de santé). Des questionnaires de suivi seront soumis régulièrement aux participants. Les modalités mises en place afin d'optimiser la collecte et l'appariement des données sont décrites.

Mots Clés : nanomatériau, collecte de données, appariement

1. Introduction

La Commission Européenne désigne sous le terme de nanomatériau (NM) tout matériau contenant des particules libres, sous forme d'agrégat ou d'agglomérat, dont au moins 50 % des particules, dans la répartition numérique par taille, présentent une ou plusieurs dimensions externes se situant entre 1 nm et 100 nm (EU, 2011), qu'il soit produit intentionnellement ou non.

Compte tenu de leurs propriétés qui apportent de la valeur ajoutée aux produits qui les intègrent, les nanomatériaux se retrouvent d'ors et déjà dans de nombreux produits disponibles sur le marché, alors même que leur toxicité potentielle n'est pas connue. En effet, les dangers potentiels pour la santé des NM ne sont pas clairement identifiés.

En France, la Direction générale de la santé et la Direction générale du travail ont confié, en 2007, à l'Institut de veille sanitaire (InVS) le développement d'un dispositif de surveillance des effets sur la santé de l'exposition professionnelle aux nanomatériaux intentionnellement produits. Une étude de faisabilité (Boutou-Kempf, 2011) a permis d'établir les modalités de surveillance à mettre en œuvre. Ce dispositif de surveillance, baptisé EpiNano repose sur la création, en France, d'un enregistrement des travailleurs potentiellement exposés, le suivi prospectif généraliste (passif et actif) de ces travailleurs auquel pourrait s'ajouter, éventuellement, une surveillance médicale et/ ou biologique. Dans un premier temps, l'enregistrement est limité aux travailleurs potentiellement exposés au dioxyde de titane nanométrique (TiO₂) et aux nanotubes de carbone (NTC) (Canu, 2013). Le but du dispositif est d'exercer un suivi généraliste à moyen et long terme des éventuels effets sur la santé de l'exposition professionnelle aux nanomatériaux manufacturés, c'est-à-dire intentionnellement produits, afin d'évaluer les risques associés à l'exposition aux NM chez l'homme. Le protocole scientifique du dispositif EpiNano a été approuvé par le Comité consultatif sur le traitement de l'information en matière de recherche dans le domaine de la santé (CCTIRS) et la Commission nationale de l'informatique et des libertés (CNIL) fin 2013 autorisant le démarrage effectif du dispositif. Les visites des entreprises et le début des inclusions des participants ont ainsi pu commencer début 2014.

L'objectif de cet article est de présenter les modalités de recueil des différentes données, ainsi que les procédés d'appariement mis en place dans le cadre du dispositif EpiNano.

¹Institut de Veille Sanitaire, 12 rue du Val d'Osne, 94415 Saint-Maurice cedex, France (epinano@invs.sante.fr)

²Université de Bordeaux-Ségalen, Laboratoire santé travail environnement (LSTE), 33076 Bordeaux, France

2. Modalités de collecte des données

2.1 Recueil des données d'exposition

Le dispositif, tel qu'il a été conçu, débute par l'identification des entreprises produisant ou utilisant des NM afin d'y évaluer la nature de l'exposition (Canu, 2013). Une fois les entreprises identifiées, une visite des installations est organisée afin d'identifier les postes potentiellement exposants, de caractériser l'exposition et de recueillir les coordonnées des personnes amenées à travailler sur ces postes. Cette visite est également l'occasion d'expliquer aux travailleurs les objectifs et la démarche du dispositif ainsi que les modalités leur permettant de l'intégrer. La qualification et la caractérisation de l'exposition sont réalisées par un binôme hygiéniste industriel/épidémiologiste de l'InVS sur la base des données qu'ils recueillent au cours de la visite.

Ces données d'exposition sont collectées via un carnet d'observation technique en entreprise (COT) (Guseva Canu, 2015(sous presse)) et vont permettre de classer les postes comme potentiellement exposants ou non. Elles vont également permettre de fournir une évaluation semi-quantitative de l'exposition.

Afin de simplifier le travail de chacun, un questionnaire en ligne, reprenant à l'identique de la structure du COT, a été développé sous un logiciel spécifique appelé Voozanoo®. Il s'agit d'un outil de conception et de création de systèmes de collecte et de gestion d'information en ligne. Les hygiénistes peuvent dorénavant saisir les données en direct depuis le lieu de la visite ou au retour de la mission, selon les possibilités techniques à leur disposition (Wifi, clé 3G, ordinateur portable, tablette numérique...). La structuration des fiches décrivant les postes de travail permet à ces dernières d'être jointes au compte-rendu de visite. Ainsi, les entreprises reçoivent, en plus du compte-rendu concluant sur le caractère potentiellement exposant ou non des postes de travail, un récapitulatif des informations relevées par les équipes de l'InVS leur permettant, s'ils le souhaitent, de mettre en place des actions pour la protection de leur personnel. De plus, le recours à un questionnaire en ligne permet la mise en place de thésaurus *a priori* sur un certain nombre de variables, évitant ainsi le classement à *posteriori* d'informations issues de zones de texte libre. Par exemple, un poste de travail sur lequel interviennent des personnes pour peser des matériaux contenant des nanotubes de carbone peut être désigné la personne en charge de la saisie par « pesée » mais également par « balance de précision ». L'introduction d'un thésaurus va permettre de désigner ce poste de manière unique sous l'intitulé « pesée ». Outre les problèmes liés à l'utilisation de majuscules, de minuscules ou d'accents, la standardisation de la saisie via un thésaurus permet à terme de simplifier le classement des données et d'éviter des étapes de regroupement d'informations, qui peuvent être génératrices d'erreurs dès lors que le volume de données saisies devient important. Des contrôles de cohérence sont mis en place sur les variables indispensables dès l'étape de saisie et les données sont directement formatées afin d'être exploitables par les logiciels statistiques, en l'occurrence SAS et Stata.

2.2 Enregistrement des salariés potentiellement exposés aux nanomatériaux

A l'issue de la visite et de l'envoi du compte-rendu, l'entreprise envoie de manière sécurisée les coordonnées des personnes travaillant à ces postes à l'InVS, qui se charge du routage des questionnaires d'inclusion.

A ce stade du dispositif, deux types de données sont recueillies : les données d'hygiène industrielle, collectées lors de la visite de l'entreprise, qui permettent de caractériser le potentiel d'exposition de chaque poste de travail et les données individuelles, recueillies via le questionnaire d'inclusion complété par les participants au dispositif. Dans l'entreprise, un même individu peut être amené à travailler sur plusieurs postes de travail. En effet, par exemple, la personne peut intervenir sur le poste de pesée pour peser la poudre contenant des nanotubes de carbone bruts de synthèse et sur le poste d'extraction où, par un procédé dit « acide », le matériau composite est dissout afin d'en extraire les nanotubes purifiés. Afin de relier ces informations entre elles, il s'est avéré nécessaire de numéroter chaque poste de travail avec un numéro unique (composé du numéro de l'entreprise, du numéro d'atelier et du numéro du poste de travail), propre à l'InVS, afin de permettre pour chaque individu d'identifier les postes de travail dans lesquels il intervient. A terme, un score individuel d'exposition, reflétant les différentes opérations effectuées à chaque poste exposant, sera calculé.

Les historiques professionnels des personnes participant au dispositif sont également recueillis, via les questionnaires d'inclusion. Chaque métier occupé est codé, après réception par des hygiénistes industriels, selon sa profession (d'après la classification des Professions et Catégories Socioprofessionnelles, version 2008) et son secteur d'activité (d'après la Nomenclature des Activités Françaises, version 2000 et 2008). Cela permet de croiser les données professionnelles issues du dispositif EpiNano avec des matrices emploi-expositions nationales (Fevotte, 2006) afin d'identifier les substances auxquels les personnes ont pu être exposées au cours

de leur carrière professionnelle. Cette information permettra d'étudier l'effet de certaines co-expositions par rapport à l'exposition aux nanomatériaux et éventuellement ajuster les modèles statistiques pour contrôler le biais de confusion ou prendre en compte un effet d'interaction entre les deux expositions. Par exemple, une exposition passée aux poussières alvéolaires de silice cristalline doit être prise en compte lorsque le lien entre exposition aux nanomatériaux et cancer broncho-pulmonaire est étudié, compte-tenu du fait que les poussières alvéolaires de silice cristalline sont connues pour être liées à un risque accru de développer un cancer broncho-pulmonaire (Pelucchi, 2006).

2.3 Suivi à moyen et long terme des participants

Le suivi à moyen et long terme se décomposera en plusieurs étapes, avec d'un côté le suivi actif, via l'envoi régulier de questionnaires de suivi et de l'autre, le suivi passif, grâce aux bases de données médico-administratives disponibles.

Concernant le suivi actif, des questionnaires de suivi seront envoyés, tous les 3 ans, afin de mettre à jour les informations relatives au statut familial, au comportement de santé (statut tabagique, pratique sportive...), au lieu de résidence, au changement de poste/emploi et au changement éventuel de comportement vis-à-vis des équipements de protection individuels. Ce questionnaire sera électronique, afin de limiter les coûts financiers (coûts d'impression, d'envoi, de saisie) tout en permettant plus de souplesse sur le contenu.

Concernant le suivi passif, il est prévu d'utiliser le fichier du Système national d'information inter-régime de l'Assurance Maladie (SNIIR-AM) pour surveiller l'apparition et la fréquence des événements de santé (protocole actuellement en cours de rédaction). Cela permettra de ne pas surcharger les questionnaires de suivi envoyés aux participants au dispositif, et de ce fait de limiter le fardeau de réponse. De plus, cela permettra de pallier aux possibles biais de mémoire (oubli, difficulté à replacer les événements dans le temps) ou de diagnostic (classement inadéquat de la maladie) des participants tout en se prémunissant du risque de non-réponse (Kone, 2009).

Les données fournies par le SNIIR-AM détaillent les remboursements de médicaments et consultations, les hospitalisations, les accidents du travail ou les maladies professionnelles (figure 2.3-1) et leur accès nécessite notamment l'utilisation du numéro d'inscription au répertoire (NIR) qui est le numéro national unique d'identification des personnes. Selon les dispositions législatives actuelles, en dehors de la sphère de protection sociale, l'utilisation du NIR nécessite un décret en Conseil d'État après avis motivé et publié de la CNIL. L'InVS a choisi de recourir à un tiers de confiance détenteur du NIR. Les données identifiantes à disposition (nom patronymique, nom de jeune fille, prénoms, sexe, date et lieu de naissance) doivent être envoyées, de manière sécurisée, à ce tiers de confiance afin que celui-ci recherche le NIR et constitue un fichier de correspondance entre le numéro d'inclusion dans le dispositif et les données nécessaires à l'extraction des données SNIIR-AM (NIR, date de naissance et sexe). Le fichier ainsi constitué peut ensuite être transmis de manière sécurisée à l'Assurance Maladie qui est en mesure d'extraire et de renvoyer à l'InVS les données de santé ou l'identifiant Assurance Maladie correspondants (figure 2.3-2). Le protocole prévoit de récupérer les informations issues du SNIIR-AM pour toute personne qui n'en aura pas manifesté le refus, c'est-à-dire pour toute personne qui n'aura pas renvoyé le formulaire de refus joint au courrier d'invitation et au questionnaire d'inclusion. Cela signifie que ces informations seront disponibles pour une partie des non-répondants au questionnaire d'inclusion. Ces données, compte tenu de leur nature et de leur diversité devraient permettre de réaliser un redressement des données collectées pour la non-réponse partielle (données manquantes chez les répondants au questionnaire d'inclusion) ou totale (non-répondants au questionnaire d'inclusion) (Santin, 2014).

Figure 2.3-1
Aperçu des informations disponibles dans le Système national d'informations inter-régime de l'assurance maladie (Sniir-am)

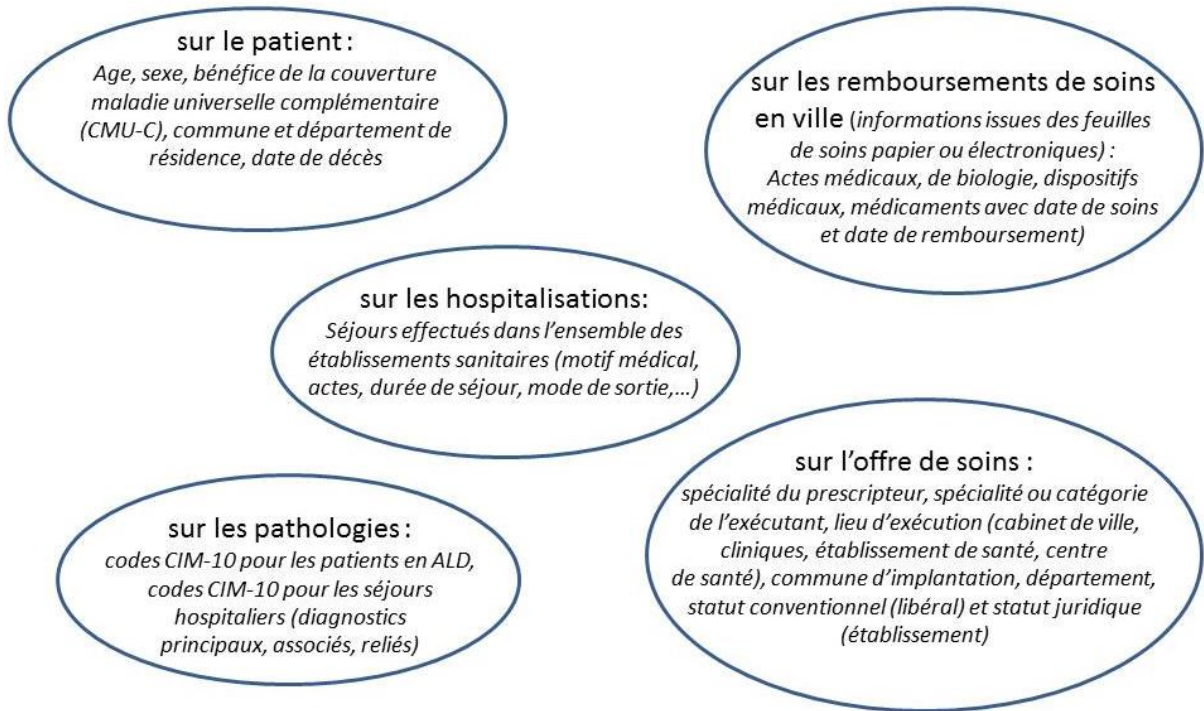
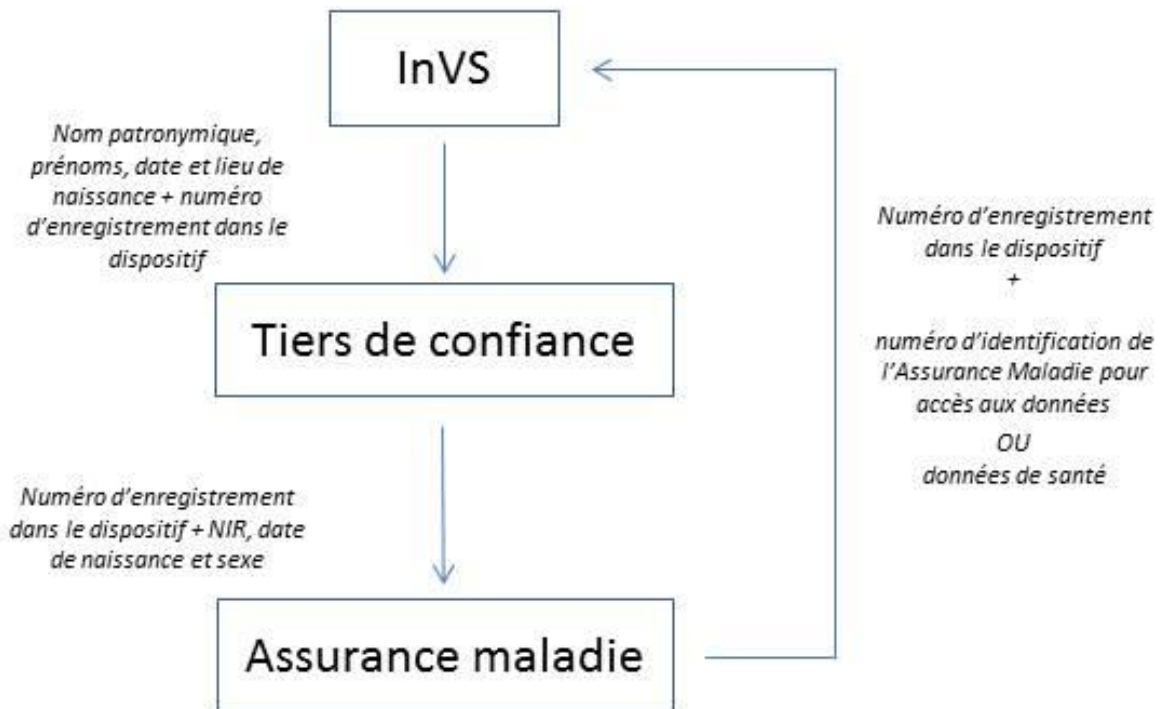


Figure 2.3-2
Dispositif EpiNano : Schéma prévisionnel de recueil des informations du Sniir-am



Outre une description de la consommation des soins médicaux des personnes, les données disponibles dans le SNIIR-AM permettent de retrouver des pathologies par le biais d'algorithme de sélection de médicaments et/ou d'actes, comme de nombreux travaux ont pu le démontrer (Moisan, 2011). En effet, certains actes traçants ou certains médicaments ont été utilisés pour identifier les sujets atteints de certaines maladies. Par exemple, en considérant l'asthme, un modèle prédisant le statut asthmatique a été construit, reposant sur des variables telles que la consultation chez un pneumologue et le remboursement de certains médicaments antiasthmatique (Iwatsubo, 2012).

Enfin, l'interrogation du Répertoire national d'identification des personnes physiques (RNIPP) et de la base des causes médicales de décès (CépiDc) permet de recueillir le statut vital des participants au dispositif, ainsi que les causes de décès le cas échéant. Pour se faire, l'InVS doit envoyer les nom patronymique (et nom de jeune fille), prénoms, sexe, date et lieu de naissance des personnes n'ayant pas exprimé le refus à l'unité U1018-Centre de recherche en Epidémiologie et Santé des Populations de l'Institut national de la santé et de la recherche médicale qui fait office de tiers de confiance lors des échanges avec le RNIPP d'une part et le CépiDc d'autre part (CépiDc, 2014).

3. Discussion

Le développement de la surveillance de l'exposition professionnelle des travailleurs français aux nanomatériaux intentionnellement produits s'inscrit dans le cadre d'une demande de la Direction générale de la santé et de la Direction générale du travail visant à mieux comprendre les risques pour la santé liés aux nanomatériaux (NM). Le schéma d'étude mis en place va permettre d'enrichir les informations individuelles issues des questionnaires d'inclusion de données d'hygiène industrielle concernant l'exposition professionnelle des participants et d'informations issues des bases de données médico-administratives allant du nombre de consultation chez le généraliste au statut vital, en passant par les médicaments remboursés ou la déclaration d'une pathologie en affection longue durée. Cela devrait permettre de répondre aux objectifs de suivi généraliste des éventuels effets sur la santé à moyen et long terme d'une exposition professionnelle aux NM, tels que prévus dans le dispositif EpiNano.

Il existe des incertitudes concernant l'organisme détenteur du NIR qui sera fera office de tiers de confiance dans le cadre des échanges avec l'Assurance Maladie. Des contacts ont été pris en ce sens auprès de la Caisse Nationale d'Assurance Vieillesse mais la collaboration entre les deux organismes n'est pas contractualisée.

Le choix de recourir, pour le suivi, à des questionnaires électroniques est principalement budgétaire. Or il peut entraîner des biais liés au mode de passation du questionnaire. En effet, l'ensemble de la population des participants n'a pas forcément accès à Internet. De plus, selon l'équipement informatique (ordinateur de bureau ou portable, tablette tactile, smartphone), le rendu visuel et les temps de chargement du questionnaire ne sont pas les mêmes, ce qui risque d'induire des biais de non-réponse qui devront par la suite être corrigés. Cela devra être mis en regard des économies réalisées.

Les modalités d'organisation du dispositif EpiNano tentent d'optimiser au mieux le recueil des données en profitant au maximum des nouvelles technologies développées ces dernières années. D'autres pistes restent à explorer, particulièrement concernant la fidélisation des participants via le développement d'une newsletter, ou d'un site web dédié par exemple.

Après une phase d'inclusion des entreprises prévue pour durer 3 ans, le suivi des participants est prévu pour une durée de 20 ans avec, à terme, une population d'étude attendue de 1 500 à 2 000 personnes. A ce jour, 10 entreprises ont été visitées, dont trois ont actuellement retourné les coordonnées de leurs employés. Cela a permis l'envoi de 54 questionnaires, avec un taux de retour de 54%.

Bibliographie

Boutou-Kempf O. (2011), *Éléments de faisabilité pour un dispositif de surveillance épidémiologique des travailleurs exposés aux nanomatériaux intentionnellement produits*, Saint-Maurice, Institut de veille sanitaire, 69 p.

Canu IG, Boutou-Kempf O, Delabre L, Ducamp S, Iwatsubo Y, Marchand JL, *et al.* (2013), « French registry of workers handling engineered nanomaterials as an instrument of integrated system for surveillance and research », *Journal of Physics : Conference Series*, 429(1)

EU. *Recommendation of European Commission of 18 October 2011 on the definition of nanomaterial.*

Févotte J, Luce D, Arslan M, Dananché B, Delabre L, Ducamp S, Garras L, Orłowski E, Pilorget C, Thuret A, Goldberg M, Imbernon E. (2006), « Surveillance des expositions d'origine professionnelle en population générale : le programme Matgéné », *Bull Epidemiol Hebd*, 46-47:362-5.

Guseva Canu I, Ducamp S, Delabre L, Audignon-Durand S, Ducros C, Durand C, Iwatsubo Y, Jezewski-Serra D, Le Bihan O, Malard S, Radauceanu A, Reynier M, Ricaud M et Witschger O. (2015), « Proposition d'une méthode de repérage des postes de travail potentiellement exposant aux nano-objets, leurs agrégats ou agglomérats dans les entreprises mettant en œuvre des nanomatériaux manufacturés », *Archives des maladies professionnelles* (sous presse).

Iwatsubo Y, Lauzeille D, Houot M, Mevel M, Chabault E, Delabre L *et al.* (2012), *Surveillance épidémiologique de l'asthme au sein de la population des artisans et commerçants affiliés au Régime social des indépendants (RSI). Résultats de l'étude pilote*, Saint-Maurice, Institut de veille sanitaire, 127 p.

Kone Pefoyo AJ, Rivard M, Laurier C. (2009), « Public health surveillance and role of administrative data », *Rev Epidemiol Sante Publique*, 57(2):99-111.

Moisan F, Gourlet V, Mazurie JL, Dupupet JL, Houssinot J, Goldberg M, *et al.* (2011), « Prediction model of Parkinson's disease based on antiparkinsonian drug claims », *Am J Epidemiol*, 174(3):354-63.

Pelucchi C, Pira E, Piolatto G, Coggiola M, Carta P, La Vecchia C. (2006), « Occupational silica exposure and lung cancer risk: a review of epidemiological studies 1996-2005 », *Ann Oncol*, 17 : 1039-50.

Santin G. *et al.* (2014), «In an occupational health surveillance study, auxiliary data from administrative health and occupational databases effectively corrected for nonresponse», *Journal of Clinical Epidemiology*, 67: 722-730

Statut vital et causes de décès – Nouvelle procédure utilisateurs d'accès eu RNIPP – Version du 19/05/2014 accessible à l'URL :

http://cesp.vjf.inserm.fr/~webifr/pdf/PROCEDURE_ACCES_RNIPP_UTILISATEURS_V2_19_05_2014.pdf