



**HAL**  
open science

## La réutilisation d'archives par les chercheurs en épidémiologie

Magalie Moysan

► **To cite this version:**

Magalie Moysan. La réutilisation d'archives par les chercheurs en épidémiologie : Analyse de la Revue d'épidémiologie et de santé publique (1968-2007). . 2015. hal-01257170

**HAL Id: hal-01257170**

**<https://hal.science/hal-01257170>**

Preprint submitted on 18 Jan 2016

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# La réutilisation d'archives par les chercheurs en épidémiologie

Analyse de la *Revue d'épidémiologie et de santé publique*  
(1968-2007)

Magalie MOYSAN

*A paraître dans La Gazette des archives en 2016.*

Depuis la naissance de l'Open Access qui promeut le libre accès aux publications scientifiques, les études sur le partage et la réutilisation des données de recherche se multiplient<sup>1</sup>. Le partage est l'un des moteurs de l'activité de recherche et n'est pas né avec Internet. Le sociologue Dominique Vinck rappelle qu'«avec les rencontres entre chercheurs, les échanges de papiers sont les opérateurs de mise en cohérence et de coordination les plus répandus. Ils participent à la construction des réseaux autant qu'ils en manifestent l'existence<sup>2</sup>». Pour autant, il faut distinguer le partage de la réutilisation. La production scientifique se nourrit des informations publiées dans des articles, mais les chercheurs réutilisent-ils des données non publiées en tant que sources ? La réutilisation soulève en effet des problèmes méthodologiques : le document a été créé à d'autres fins et son contexte de production n'est pas toujours maîtrisé.

---

<sup>1</sup> Pour un panorama des études sur le sujet, voir WALLIS (Jillian), ROLANDO (Elizabeth), BORGMAN Christine, « If We Share Data, Will Anyone Use Them? Data Sharing and Reuse in the Long Tail of Science and Technology », *PLoS One*, n° 8, juillet 2013. DOI: 10.1371/journal.pone.0067332

<sup>2</sup> VINCK (Dominique), « Les objets intermédiaires dans les réseaux de coopération scientifique. Contribution à la prise en compte des objets dans les dynamiques sociales », *Revue française de sociologie*, n° 40, 1999, p. 405.

Dans cet article, nous nous intéresserons aux conditions et pratiques de réutilisation dans une discipline médicale, l'épidémiologie, en nous concentrant sur les documents et données d'archives : quelles archives sont utilisées par les chercheurs et sous quelle terminologie ? Comment s'intègrent-elles dans le processus de recherche ? Pour répondre à ces questions, nous avons dépouillé les articles publiés dans la *Revue d'épidémiologie et de santé publique* de 1968 à 2007.

### **Méthodologie adoptée**

L'épidémiologie est une discipline de la recherche médicale qui vise à réaliser « une étude de la distribution des maladies et des invalidités dans les populations humaines, ainsi que des influences qui déterminent cette distribution<sup>3</sup> ». Elle offre des caractéristiques intéressantes pour une étude archivistique. Contrairement aux autres spécialités médicales, elle se concentre sur la santé des populations et non des individus. Elle a recours à des sources variées et compile de grands volumes de données grâce à des méthodes statistiques. Par ailleurs, elle comporte une forte dimension préventive et privilégie une approche sur le temps long qui favorise l'émergence de tendances.

En France, les travaux de recherche sont notamment publiés dans la *Revue d'épidémiologie et de santé publique* coordonnée par l'Institut national de la santé et de la recherche médicale (INSERM). Créée en 1953 sous le titre *Revue d'hygiène et de médecine sociale*, elle devient *Revue d'épidémiologie, médecine sociale et santé publique* en 1971 et adopte sa dénomination actuelle en 1976. Cette publication répond à un besoin de faire connaître l'actualité dans le domaine de l'hygiène puis de l'épidémiologie, de réunir une communauté éparpillée (acteurs du secteur public et privé, armées, hôpitaux, ...) et de mettre en valeur des travaux récents<sup>4</sup>. Pour notre étude, elle présente l'intérêt de paraître à intervalles

---

<sup>3</sup> Définition de l'Organisation mondiale de la Santé (1968).

<sup>4</sup> PARISOT (Jacques), « Avant-propos », *Revue d'hygiène et de médecine sociale*, n° 1, 1953, p. 5.

réguliers depuis plus de quarante ans, dans un format qui a peu évolué. Et surtout, dès ses débuts, tous les articles originaux sont structurés en deux grandes parties, « matériel et méthodes » (ou « méthodes ») et « résultats ». Dans la première partie, les chercheurs décrivent leurs sources et leur méthodologie. Cette rubrique, par son caractère systématique, constitue un terrain privilégié pour étudier les pratiques de réutilisation. Nous avons donc dépouillé systématiquement l'ensemble des articles de recherche publiés dans la Revue entre 1968 et 2007, en consignant tous les cas d'utilisation de documents ou de données d'archives.

La difficulté majeure était de proposer une définition des archives, alors même que le terme n'est pas couramment employé par les chercheurs. Selon le Code du patrimoine, « les archives sont l'ensemble des documents produits et reçus, quels que soient leur date, leur lieu de conservation, leur forme et leur support, produits ou reçus par toute personne physique ou morale et par tout service ou organisme public ou privé dans l'exercice de leur activité<sup>5</sup> ». Cette définition nous a servi de base pour établir ce que nous entendons par « archives » :

- Les archives ne sont pas uniquement des informations, mais des documents et des données, composés d'un support et d'un contenu.
- Elles se distinguent de la documentation par leur organisation selon le principe de provenance.
- Les archives s'inscrivent dans une dimension temporelle et renvoient au passé, récent ou plus ancien. Pour notre étude, nous distinguons deux moments : le moment de la création et de l'utilisation première, et le moment de la réutilisation. Ces deux moments ne sont pas toujours strictement successifs.
- Le document ou la donnée deviennent archives par un processus de conservation ou de « mise en archive », qui intervient dès leur création ou lorsqu'ils n'ont plus d'utilité immédiate.
- Ce processus de conservation induit un changement de statut, de finalité et d'usage pour les documents ou les données, qui sont alors utilisés à d'autres fins que celles qui ont présidé à leur création. Notre corpus est ainsi constitué :

---

<sup>5</sup> Code du patrimoine, article L 211-1.

- de sources détournées de leur finalité initiale à des fins de recherche (registres d'état civil, dossiers médicaux de patients) ;
  - des sources produites par la communauté scientifique sans projet de recherche précis (bulletins de décès) ;
  - d'archives scientifiques réutilisées à d'autres fins que la recherche initiale.
- Enfin, nous excluons des archives les publications scientifiques qui obéissent à des logiques de partage et de conservation spécifiques.

Nous considérons donc que les archives sont des documents et données issus d'un producteur, qui ont fait l'objet d'un processus de conservation dès leur création ou à l'issue d'une première utilisation. Sur cette base, nous avons consigné l'ensemble des articles de recherche médicale publiés dans la *Revue d'épidémiologie et de santé publique* qui portent des traces de réutilisation d'archives. Notre périmètre s'est limité à la France et seuls les articles co-signés par un ou des chercheurs qui travaillent sur le territoire français ont été intégrés à notre étude.

Le dépouillement a abouti à la constitution d'un corpus de 330 articles qui font état d'une réutilisation d'archives entre 1968 et 2007. Ces publications représentent 21% des articles originaux parus dans la Revue, ce qui nous permet de tirer un premier enseignement : si les pratiques de réutilisation existent, elles ne sont pas la norme pour les chercheurs qui publient dans la Revue. Leurs analyses sont prioritairement fondées sur des observations cliniques ou des questionnaires collectés à des fins immédiates de recherche.

## **Quelles sources d'archives pour la recherche ?**

La réutilisation d'archives implique un détournement ou une réorientation par rapport à leur usage initial. Mais quel était-il ? Nous nous sommes intéressés dans un premier temps à la provenance des archives réutilisées.

Le dépouillement des articles de notre corpus montre que les chercheurs de la *Revue d'épidémiologie et de santé publique* privilégient l'univers médical. Près de 36% des articles font usage de sources provenant des hôpitaux. Plusieurs raisons pourraient l'expliquer : les hôpitaux génèrent de grands volumes de données, centralisées et standardisées par le Programme médicalisé des systèmes d'information (PMSI) à partir de 1991 ; leur activité se concentre sur des pathologies plus graves et plus complexes que les médecins généralistes et intéresse davantage la recherche médicale ; par ailleurs, les chercheurs exercent en majorité leurs activités dans des laboratoires installés à proximité des centres hospitaliers. Ils ont donc un accès privilégié à ces informations. A l'inverse, seuls 8,6% des articles font usage de dossiers conservés par des médecins non hospitaliers. 16% des articles retenus exploitent quant à eux des données issues des administrations de santé comme le Ministère de la Santé, les caisses primaires d'assurance maladie ou les mutuelles. Ces organismes prêtent régulièrement leur concours à la recherche médicale en donnant accès à leurs bases de données<sup>6</sup>.

La recherche médicale produit également des sources, telles que les registres de morbidité utilisés dans une quarantaine d'articles. Le registre est « une structure épidémiologique qui réalise l'enregistrement continu et exhaustif des cas d'une pathologie donnée, dans une région géographique donnée, et qui, à partir de cet enregistrement, réalise seule ou en collaboration, des études visant à améliorer les connaissances concernant cette pathologie<sup>7</sup> ». Il recense donc tous les cas d'une pathologie sur un territoire, par exemple tous les cancers du département du Bas-Rhin. Créés à partir des années 1980 pour les besoins de la recherche, les registres constituent une source privilégiée pour l'étude de certaines maladies<sup>8</sup>. Les bulletins et statistiques de décès sont également contrôlés par la recherche médicale et servent aux nombreuses études sur les

---

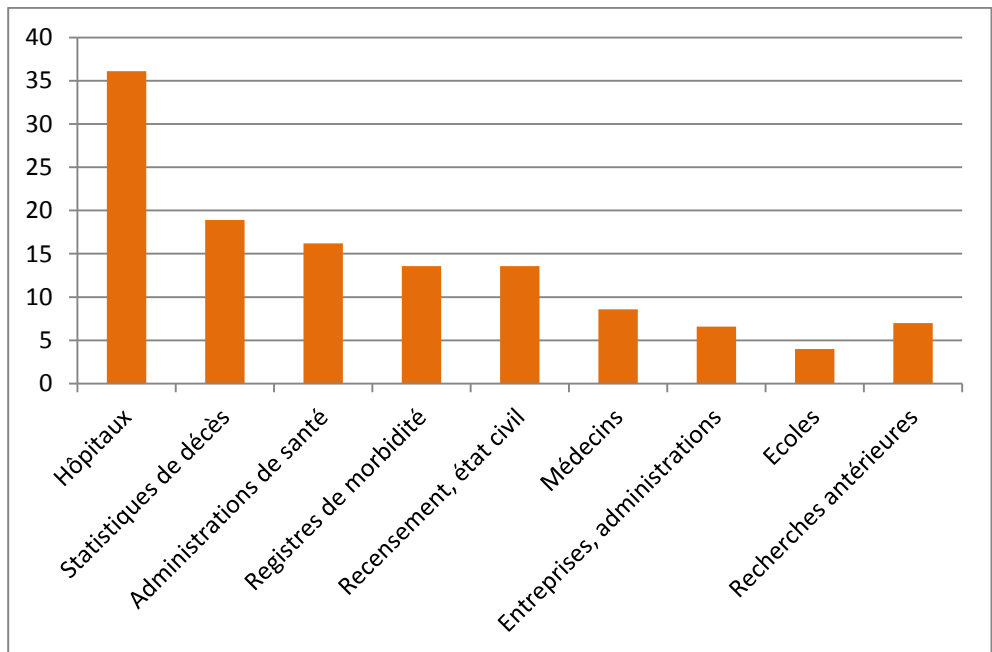
<sup>6</sup> Voir par exemple ZMIROU (Denis), PARENT (Benjamin), POTELON (Jean-François), « Etude épidémiologique des effets sur la santé des rejets atmosphériques d'une usine d'incinération de déchets industriels et ménagers », *Revue d'épidémiologie et de santé publique*, n° 32, 1984, p. 393.

<sup>7</sup> « Eléments pour une politique nationale des registres », *Bulletin épidémiologique hebdomadaire*, n° 5, 1989, p. 17.

<sup>8</sup> DUCIMETIERE (Pierre), MONTAVILLE (Bernard), SCHAFFER (Paul), SPIRA (Nadine) sous dir., *Recherche et politiques de santé. L'apport des registres de morbidité*, Paris, La documentation française, 1992, p. 1.

causes de mortalité. Un article sur cinq y fait référence dans notre corpus. Chaque décès constaté par un médecin est décrit dans une fiche envoyée à l'INSERM, qui l'intègre dans une base de données. Cette dernière est ensuite mise à la disposition des chercheurs qui en font la demande.

Enfin, l'état civil et les informations de recensement sont consultés dans 13,6% des articles. Ils servent principalement de sources secondaires, pour connaître le statut d'un individu (décédé, marié, *etc.*) ou confronter les résultats scientifiques obtenus à la population d'un territoire.



*Origine des sources utilisées par les chercheurs*

Les épidémiologistes privilégient logiquement les sources d'origine médicale, de préférence faciles d'accès. Les informations utilisées sont souvent produites pour la recherche (par exemple les registres) ou sous une forme standardisée qui facilite la réutilisation (statistiques de décès, données hospitalières).

## Des archives scientifiques peu réutilisées

En revanche, les épidémiologistes réutilisent peu d'archives issues de projets de recherche antérieurs : seuls 21 articles y font mention. Dans plus de la moitié des cas, le chercheur réutilise des informations à la création desquelles il a contribué. Dans les autres cas, il fait usage de données produites localement (peut-être connaît-il le producteur des archives ?) ou issues d'une étude à notoriété nationale.

Plusieurs hypothèses peuvent être envisagées pour expliquer cette faible réutilisation. La première serait que le chercheur se réapproprie très peu les matériaux issus de projets précédents. Une étude menée au Center for Embedded Network Sensing (CENS), un centre de recherche pluridisciplinaire de l'Université de Californie Los Angeles, en arrive à cette conclusion. Réalisée entre 2002 à 2012 à partir d'entretiens et d'observations de terrain, elle montre que les chercheurs travaillent souvent en relation avec les observatoires et organismes publics chargés de collecter des données. Mais seule une minorité d'entre eux reconnaît utiliser des sources reçues par des confrères, en mettant en avant leur manque de confiance dans le processus de collecte. Lorsqu'ils réutilisent la production scientifique d'un collègue, ils privilégient les informations recueillies dans un cercle restreint et de confiance. Par ailleurs, ces sources viennent généralement en appui de celles qu'ils ont eux-mêmes collectées et sont rarement fondamentales pour le projet de recherche<sup>9</sup>.

Nous pourrions à l'inverse supposer que la réutilisation d'archives scientifiques n'est pas mentionnée car elles sont utilisées en appui des publications déjà citées. Ainsi, les chercheurs ne distingueraient pas la citation des publications et celle des données. L'analyse des articles de la revue ne suffit pas à privilégier

---

<sup>9</sup> WALLIS (Jillian), ROLANDO (Elizabeth), BORGMAN (Christine), « If We Share Data, Will Anyone Use Them? Data Sharing and Reuse in the Long Tail of Science and Technology », *Plos One*, n°8, juillet 2013, p. 6-15. Disponible sur : <http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0067332>



l'une ou l'autre de ces hypothèses : il conviendrait de mener des entretiens avec des épidémiologistes pour en savoir davantage.

## **Les mots et les choses : les archives dans le vocabulaire des chercheurs**

Dans le processus de création scientifique, les chercheurs utilisent des sources que nous assimilons à des archives. Mais comment les nomment-ils ? Pour le savoir, nous avons sélectionné dix termes susceptibles d'être utilisés par les chercheurs pour décrire ces sources, tels que « document », « archives », « dossiers » ou « données ». Nous avons ensuite noté l'apparition d'un ou plusieurs de ces termes dans la partie « matériel et méthodes » de chaque article.



*Termes utilisés dans la Revue d'épidémiologie et de santé publique pour caractériser des archives*

Comme le montre le nuage de mots ci-dessus, les chercheurs qualifient principalement les archives réutilisées de « données », que l'on retrouve dans près de 78% des articles. D'après le dictionnaire de l'Académie française, le terme existe depuis le 18<sup>e</sup> siècle sous cette signification. Il est en effet déjà employé dans 31% des articles publiés entre 1968 et 1979. Son utilisation explose à partir des années 1980, peut-être popularisée par l'adoption de la loi Informatique et Liberté de 1978, qui impose la déclaration des traitements de données à caractère personnel, très courants dans la recherche médicale.

Le deuxième terme utilisé par les épidémiologistes est « informations », que l'on retrouve dans 34,6% des articles. Tout comme « données », il semble privilégier une approche « dématérialisée » de la source et met en valeur le contenu davantage que le support. En troisième position, « dossier » renvoie quant à lui clairement à la matérialité des sources, mais il sert principalement à désigner le dossier médical.

En revanche, les archives apparaissent assez peu dans le vocabulaire des chercheurs. Nous les retrouvons seulement dans 17 articles, qui n'offrent qu'un aperçu de ce qu'ils associent à ce terme. Un article oppose le dossier patient aux archives du service et donne l'impression que les archives représentent l'ensemble des documents non sériels du producteur. Par ailleurs, dans ce corpus de 17 articles, les auteurs identifient plus facilement les documents produits par une entreprise ou une administration comme des archives. Peut-être cette dénomination est-elle liée à la position des documents chez le producteur (local séparé, *etc.*) ? Mais là encore, seule une étude complémentaire fondée sur des entretiens permettrait d'aller plus loin.

## Des archives parfois anciennes

Nous avons porté notre intérêt sur l'épidémiologie, car nous supposons que son objet d'étude – la distribution des maladies dans les populations – conduisait les scientifiques à mener des enquêtes sur le long terme.

L'analyse de notre corpus confirme cette hypothèse. Nous avons étudié l'écart entre la date de création des sources utilisées et la date de publication d'un article. Près de 38% des articles réutilisent des archives qui ont moins de cinq ans au moment de la publication, parfois créées très récemment avant leur réutilisation. Mais les chercheurs consultent également des sources plus anciennes, produites cinq à dix ans auparavant (28% des articles), voire entre dix et quinze ans avant la publication (16,5%). Enfin, 17,5% des articles du corpus utilisent des sources de plus de quinze ans. Certains projets de recherche couvrent une très large période : l'étude des certificats de décès depuis 1925<sup>10</sup>, l'analyse des archives d'un centre de rééducation de 1949 à 1997<sup>11</sup> ... Les chercheurs semblent donc prêts à recueillir des informations très anciennes pour les besoins de leurs recherches.

Pour faciliter l'observation d'un phénomène sur la longue durée, l'épidémiologie a également constitué des cohortes, « ensembles de sujets ayant vécu une même expérience, et suivis dans le temps depuis la date de cette expérience<sup>12</sup> ». L'équipe scientifique sélectionne une catégorie de population (habitants d'une ville, personnes âgées, *etc.*) correspondant à ses thématiques de recherche (maladies cardiovasculaires, vieillissement, *etc.*). Les personnes retenues sont régulièrement sollicitées pour subir des examens médicaux ou

---

<sup>10</sup> CADET (B.), ROUBINE (J.-M.) LEIBOVICI (Didier), « Le passé et le présent de la mortalité asthmatique : vers une rupture de la tendance séculaire? », *Revue d'épidémiologie et de santé publique*, n° 41, 1993, p. 363-373.

<sup>11</sup> LHERITIER (K.), RAVAUD (Jean-François), DESERT (J.F.), PEDELUCQ (J.P.), O'HANNA (F.), DAURES (Jean-Pierre), « La survie des blessés médullaires tétraplégiques après la première entrée en centre de rééducation et ses facteurs pronostiques », *Revue d'épidémiologie et de santé publique*, n° 49, 2001, p. 449-458.

<sup>12</sup> LECLERC (Annette), PAPOZ (Laure), BRÉART (Gérard), LELLOUCH (Joseph), *Dictionnaire d'épidémiologie*, Paris, Frison-Roche, 1990, p. 26.

répondre à des questionnaires. Les chercheurs se constituent ainsi des bases de données qui couvrent une large période. Les cohortes les plus anciennes ont été créées après la Seconde Guerre mondiale. C'est le cas de l'enquête de Framingham qui réalise depuis 1948 un suivi médical des habitants d'une ville moyenne pour établir les relations entre génétique et maladies cardiovasculaires<sup>13</sup>. En France, quelques cohortes font figure de référence : l'enquête Paquid qui étudie les phénomènes de démence depuis 1988<sup>14</sup> ou encore le projet GAZEL sur la santé des salariés d'Electricité de France-Gaz de France conduit depuis 1989<sup>15</sup>.

### **Les archives dans le processus de recherche : le document, l'information et la donnée**

La réutilisation de sources anciennes implique la mise en place de procédures de conservation. L'archivage est principalement assuré par les producteurs des documents et données : hôpitaux, administrations de santé, mairies, entreprises, *etc.*

Mais la recherche médicale s'est aussi organisée pour préserver les sources dont elle avait besoin et a mis en place des systèmes de conservation centrés sur l'information. Ce n'est pas l'archivage du document au sens classique du terme qui intéresse les chercheurs, mais son codage et son enregistrement. Qu'il

---

<sup>13</sup> GIROUX (Emilie), « Origine de l'étude prospective de cohorte », *Revue d'histoire des sciences*, n° 64-2, 2011, p. 304-308.

<sup>14</sup> DARTIGUES (Jean-François), HELMER (Catherine), LETENNEUR (Luc), PERES (Karine), AMIEVA (Hélène), AURIACOMBE (Sophie), ORGOGOZO (Jean-Marc), COMMENGES (Daniel), JACQMIN-GADDA (Hélène), RICHARD-HARSTON (Sandrine), DELVA (Fleur), FOUBERT-SAMIER (Alexandra), BARBEGER-GATEAU (Pascale), « Paquid 2012 : illustration et bilan », *Gériatrie et psychologie Neuropsychiatrie du vieillissement*, n° 10, 2012, p. 325-331. DOI : [10.1684/pnv.2012.0353](https://doi.org/10.1684/pnv.2012.0353)

<sup>15</sup> Sur l'histoire de la cohorte GAZEL, voir ZINS (Marie), LECLERC (Annette), GOLDBERG (Marcel), « The French GAZEL Cohort Study : 20 years of epidemiologic research », *Advances in Life Course Research*, n° 14, 2009, p. 135-146. DOI : [10.1016/j.alcr.2010.02.003](https://doi.org/10.1016/j.alcr.2010.02.003)

s'agisse des bulletins de décès, des cohortes ou des registres de morbidité, cette procédure prend souvent la même forme, très bien décrite dans le cas des registres.

La constitution d'un registre se fait en plusieurs étapes. Les chercheurs commencent généralement par analyser les informations contenues dans les archives sur une pathologie. Un article du registre des tumeurs digestives du Calvados décrit le processus de collecte :

« De 1978 à 1984, tous les nouveaux cas de cancers colorectaux ont été enregistrés de manière active par le Registre spécialisé des tumeurs digestives du Calvados, département peuplé de 589.260 habitants selon le recensement de 1982. Le recueil des données a été effectué à partir des données des laboratoires d'anatomopathologie publics et privés, des dossiers d'hospitalisation des établissements de soin, des dossiers des spécialistes gastro-entérologues et complété par des sources administratives. Cependant, les données connues par les seules sources administratives n'étaient pas prises en compte non plus que les données purement autopsiques<sup>16</sup> ».

Une fois le registre mis en place, les informations sont systématiquement envoyées aux coordinateurs du registre par les médecins partenaires, qui soumettent leurs dossiers à une procédure d'enregistrement et de codification. Le document exploité devient information pour faciliter son utilisation à des fins de recherche. Ces informations sont ensuite consignées dans une base de données qui servira de support aux réutilisations futures. Et les chercheurs ne reviennent au document original que lorsqu'ils réalisent des tests de fiabilisation ou font évoluer leurs procédures.

Dès ses débuts, l'épidémiologie cherche à homogénéiser et fiabiliser l'information par le codage. Plusieurs articles de la *Revue d'épidémiologie et de santé publique* font état de la préoccupation des chercheurs pour établir une

---

<sup>16</sup> POTTIER (Didier), LAUNOY (Guy), LE CONTOUR (Xavier), DUGLEUX (Georges), GIGNOUX (Marc), « Importance de la définition des zones urbaines dans l'étude des facteurs de risque de cancers », *Revue d'épidémiologie et de santé publique*, n° 37, 1989, p. 392.

classification fine des informations<sup>17</sup> et améliorer la fiabilité des données collectées<sup>18</sup>. Ainsi, un papier publié en 2006 souligne que le dossier patient présente souvent un manque d'information et donc de fiabilité et lui préfère les grosses bases de données, jugées plus pérennes<sup>19</sup>.

Par les procédures de codage, l'épidémiologie a recours à des méthodes statistiques qui remontent au 19<sup>e</sup> siècle et à la naissance de l'Etat-nation selon Alain Desrosières : « A ce moment, la statistique passe du manuscrit enfermé dans les archives de l'administration, à l'imprimé destiné en principe à un large public<sup>20</sup> ». Le codage n'est donc pas né avec l'informatique et, dès ses débuts, l'épidémiologie moderne a théorisé et adapté ces méthodes à ses spécificités.

Pour la recherche épidémiologique, le document original importe peu, car les besoins s'expriment en termes de fiabilité et non de preuve. Si l'information a été collectée selon une méthode fiable, elle peut être réutilisée sans revenir au document original. Le document est réduit à son contenu informationnel conservé ensuite dans des bases de données, sur lesquelles semble se déporter la préoccupation de conservation sur le long terme.

Le rôle de l'archiviste n'est jamais mentionné dans la *Revue d'épidémiologie et de santé publique* et les chercheurs fréquentent apparemment peu les services d'archives non hospitaliers. Les documents et bases de données produits sont organisés selon les besoins de la recherche médicale et ne suivent pas les principes archivistiques fondamentaux tels que le respect des fonds. Pour autant, les études épidémiologiques qui couvrent une large période tendent à montrer que la discipline a réussi le pari d'une conservation à long terme des informations dont elle avait besoin.

---

<sup>17</sup> LECLERC (Annette), « Utilisation de la classification préalablement au tirage d'un échantillon », *Revue d'épidémiologie et de santé publique*, 1974, n° 22, p. 153-167.

<sup>18</sup> DUSSAUCY (A.), VIEL (J.F.), MULIN (B.), EUVRARD (J.), « L'outil PMSI : biais, sources d'erreurs et conséquences », *Revue d'épidémiologie et de santé publique*, 1994, n°42, p. 345-358.

<sup>19</sup> TOUZET (S.), SCHOTT (A.-M.), PLOIN (D.), MAUJEAN (G.), COLIN (C.), « L'évaluation des pratiques cliniques : quels modes de recueil des données et à partir de quelles sources d'information ? », *Revue d'épidémiologie et de santé publique*, 2006, n° 54, 223-232.

<sup>20</sup> DESROSIERES (Alain), *La politique des grands nombres. Histoire de la raison statistique*, Paris, La Découverte, 2000 (Première édition : 1993), p. 47.

Magalie MOYSAN

Doctorante en archivistique à l'université d'Angers et archiviste à l'université Paris Diderot

[Magalie.moysan@gmail.com](mailto:Magalie.moysan@gmail.com)

**La réutilisation d'archives par les chercheurs en épidémiologie. Analyse de la *Revue d'épidémiologie et de santé publique* (1968-2007)**

Cet article porte sur les pratiques de réutilisation d'archives en épidémiologie, analysées à partir d'un dépouillement de la *Revue d'épidémiologie et de santé publique* entre 1968 et 2007. Nous assimilons ici à des archives les documents et données qui ont fait l'objet d'un processus de conservation dès leur création ou à l'issue d'une première utilisation.

Les épidémiologistes réutilisent majoritairement des archives de santé, produites par les hôpitaux (dossiers de patients, *etc.*), par les administrations de santé telles que les mutuelles, ou par la recherche médicale elle-même. Les archives réutilisées sont parfois très anciennes. La discipline s'est donc dotée de systèmes de conservation qui privilégient l'information au document. Par des procédures d'enregistrement et de codage, elle conserve sur le long terme des bases de données qui répondent aux besoins de recherche.