

# Aspects éthiques et valorisation scientifique des cohortes

**Divers aspects relevant de l'éthique de la recherche sont présents dans les études effectuées sur des cohortes.**

*Les références entre crochets renvoient à la Bibliographie générale p. 51.*

**Anne Cambon-Thomsen  
Laurence Mabile  
Emmanuelle Rial-Sebbag**

UMR U 1027, Inserm, Épidémiologie et analyses en santé publique : risques, maladies chroniques et handicaps, faculté de médecine, université de Toulouse III Paul Sabatier, UMR 1027, Département d'épidémiologie et de santé publique

Les cohortes relèvent de la recherche sur l'homme et mettent en œuvre des recrutements de personnes, en bonne santé ou malades. Le recueil d'informations les concernant comporte des données nominatives, par questionnaire et/ou accès à des dossiers de santé, souvent des collections d'échantillons biologiques qui permettent de générer à leur tour de nouvelles informations personnelles, notamment s'il s'agit de données génétiques considérées comme sensibles. De plus, le suivi longitudinal des personnes, la communication d'informations issues des recherches, outre leur publication scientifique, soit individuellement aux participants, soit en tant que résultats généraux vers le public et les professionnels, sont aussi des questions aux multiples facettes. Certains aspects sont toutefois spécifiques aux cohortes. Ainsi, la réalisation des cohortes, des bases de données associées, des biobanques qui les accompagnent, l'échange de ces données et échantillons qui conditionnent nombre de recherches aujourd'hui ont fait apparaître plus récemment la nécessité de la reconnaissance des efforts requis pour ce travail essentiel. Nous traiterons donc des enjeux éthiques :

- des cohortes dans le cadre général de la recherche biomédicale et de leur traduction en pratique, notamment telle qu'encadrée spécifiquement par la loi ;
- spécifiques aux cohortes actuelles, et liés notamment à la taille et à la durée des études ainsi qu'aux technologies utilisées ;
- liés à l'évolution de la société et à l'implication plus grande dans la recherche des participants ou de leurs associations ;
- qui s'attachent à l'utilisation des cohortes

par de nombreuses équipes et à la reconnaissance du travail spécifique à cette mise à disposition. Les obstacles au partage, les mesures incitatives et certains outils possibles seront brièvement exposés.

## **Les aspects éthiques des cohortes dans le cadre général de la recherche biomédicale**

Les aspects éthiques sont en France régulés pour la plupart par des dispositions légales [38]. Ils relèvent du respect de *l'autonomie des personnes*, traduit principalement par *une information et un consentement individuel* adéquats ; cela est réalisé selon les modalités indiquées dans le Code de la santé publique dans les articles relatifs aux recherches qui comportent notamment l'examen du protocole par un Comité de protection des personnes (CPP) et, le cas échéant, son enregistrement auprès de l'Afssaps (Agence française de sécurité sanitaire des produits de santé). La récente loi n° 2012-300 du 5 mars 2012 relative aux recherches impliquant la personne humaine concerne non seulement les recherches biomédicales classiques, mais étend son champ d'application en l'aménageant aux recherches non interventionnelles, parmi lesquelles les recherches épidémiologiques.

Le principe de la *confidentialité des informations recueillies et du respect de la vie privée* est mis en œuvre en particulier par des mesures d'anonymisation et de protection de l'accès aux données, soumises à l'autorisation de la Commission nationale informatique et libertés (Cnil). Cette commission a été créée par la loi du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés,



qui a été complétée par l'article 42-2 de la loi n° 94-548 du 1<sup>er</sup> juillet 1994 relative au traitement de données nominatives ayant pour fin la recherche dans le domaine de la santé, et modifiée par la loi du 6 août 2004.

Ces aspects éthiques relèvent également du respect de la *proportionnalité* entre le type des données conservées, la durée de cette conservation, les méthodes prévues pour l'analyse, la protection de la vie privée et les besoins de la recherche dont l'appréciation est faite par un comité particulier, le CCTIRS (Comité consultatif sur le traitement de l'information en matière de recherche dans le domaine de la santé), créé par un décret paru en 1995 et qui doit être saisi pour avis préalablement à la Cnil.

Il existe des *mesures relatives aux collections d'échantillons biologiques* utilisées pour la recherche qui comportent une déclaration au ministre chargé de la Recherche ou une autorisation selon les cas, et notamment en cas de cession de la collection, visant, dans le respect des personnes, à assurer la valorisation et l'utilisation des collections. Ces procédures sont mentionnées aux articles L. 1243-3 et L. 1243-4 du Code de la santé publique. Lorsque l'organisme est un établissement de santé, la déclaration est faite conjointement au ministre chargé de la Recherche et au directeur de l'agence régionale de santé territorialement compétent. La déclaration est soumise concomitamment pour avis au Comité consultatif de protection des personnes (CPP, cité plus haut). L'utilisation secondaire en recherche des éléments du corps humain à une autre fin que celle initialement décrite est possible et soumise à des mesures particulières :

- soit information et non-opposition, selon l'article L. 1211-2 du Code de la santé publique, démarche moins contraignante qu'un nouveau consentement explicite ;

- soit, s'il est impossible de délivrer cette information ou si le CPP ne considère pas cette démarche nécessaire et donne un avis favorable à cette utilisation pour une autre finalité. La récente loi n° 2012-300 du 5 mars 2012 relative aux recherches impliquant la personne humaine ouvre sous certaines conditions la non-opposition aux utilisations secondaires de recherche génétique.

Enfin il existe des mesures d'*autorisation d'importation et d'exportation* des échantillons biologiques en cas de collaboration internationale, en application des articles R. 1235-7 et R. 1235-8 du Code de la santé publique.

Ces aspects encadrés par la loi concernent les cohortes au sein des recherches impliquant la personne humaine mais ne recouvrent

pas l'ensemble des enjeux éthiques liés à la constitution et à l'utilisation des cohortes aujourd'hui.

### Enjeux éthiques spécifiques aux cohortes actuelles, liés à l'évolution de la recherche

Ils sont de divers ordres.

La taille et la durée des études longitudinales, alliées aux technologies utilisées, amènent à constituer au fil des années de très grandes bases de données personnelles qui doivent être accessibles à la communauté scientifique. Cette ouverture est la condition même de leur mise en œuvre et de leur exploitation scientifique. Cela a été rappelé récemment à l'occasion de la parution de la Charte sur les modalités d'ouverture à la communauté scientifique des cohortes en santé [7], qui sont considérées, plutôt que comme des projets de recherche en elles-mêmes, comme des infrastructures de recherche, sur lesquelles se branchent de nombreux projets pendant une longue période. L'accumulation en leur sein de données biologiques issues des analyses effectuées par des technologies à haut débit ; par exemple, des études génomiques sur génome entier, qui étudient non pas un ou un petit groupe de gènes, mais une très large part du matériel génétique, et qui commencent à être réalisées sur des cohortes de grande taille interrogent les cadres éthiques classiques [4]. La notion d'anonymisation de telles données par exemple est mise à mal et le croisement de jeux de données qui, pris individuellement, respectent l'anonymat peut amener à des ré-identifications. Une tension peut exister entre nécessité de partager ces données et de les ouvrir à une large communauté scientifique et protection des personnes.

Les études sur le long terme nécessitent une gestion particulière des consentements. Ainsi, sur des cohortes d'enfants, il faudra prévoir leur assentiment, ainsi qu'une information adaptée à leur degré de maturité, puis leur consentement lorsqu'ils atteignent l'âge adulte. Chez les personnes âgées, c'est au contraire la perte possible de la capacité à consentir au cours du suivi qu'il faut prévoir.

La communication de résultats avec la mise à jour de facteurs, notamment biomarqueurs, prédictifs d'états de santé pose également de nombreuses questions. Quel degré de validation est-il nécessaire avant de pouvoir rapporter des résultats individuels aux participants qui le désirent ? Selon quelles modalités ? Que faire des marqueurs prédictifs dans les cohortes d'enfants ? Comment gérer le droit

de savoir ou de ne pas savoir en face de la complexité des données ?

Générer des jeux de données pour lesquelles les capacités d'interprétation vont évoluer avec le temps et l'évolution des connaissances amène donc à réfléchir sur la prise en compte de nouvelles questions éthiques au fil du déroulement des projets de grandes cohortes. En effet, un avis d'un comité d'éthique au début du projet, bien qu'évidemment nécessaire, est loin d'être suffisant pour faire face aux nombreuses interrogations qui vont apparaître pendant la constitution et l'exploitation d'une cohorte. Il faut donc prévoir, au sein même du protocole de cohorte, un système de gouvernance éthique robuste, avec parfois mise en place de comités ou groupes spécifiques à ces cohortes.

Partager les données est au cœur des enjeux éthiques, au-delà de la protection des personnes : des tensions peuvent apparaître entre la nécessité d'ouverture des cohortes à la communauté scientifique, de plus en plus internationale, et le temps d'exclusivité d'exploitation des données pour les personnes ou les équipes qui ont œuvré pour mettre en place une cohorte, au vu des capacités d'évaluation et de reconnaissance de ce travail et de l'organisation de ce partage. Cela constitue effectivement un pan entier de la recherche, organisé différemment suivant les communautés scientifiques [32]. Cette dimension est souvent mal identifiée et mal reconnue. Les obstacles au partage sont donc nombreux et des outils techniques et de management aussi bien que des cadres éthiques sont nécessaires dans ce domaine.

Nous avons donc vu que, à côté de l'encadrement éthique et légal classique en recherche biomédicale et en épidémiologie, l'évolution scientifique et la façon dont se fait la recherche aujourd'hui interrogent largement ce cadre et amènent à considérer de nouvelles dimensions. Cependant cette évolution n'est pas isolée et les relations science/société évoluent aussi.

### Les aspects éthiques liés à l'évolution de la société

Les relations science/société ou l'ancrage de la recherche dans des questions de société sont particulièrement prégnants dans le cadre des cohortes en santé. Bien sûr, du fait que de telles infrastructures reposent sur la participation de personnes, leur motivation soutenue et à long terme. Mais aussi parce que l'implication même des personnes dans la recherche a évolué ces dernières années. Passer de « objet de recherche » à « sujet de recherche » marquant

déjà une évolution, concrétisée par l'installation du cadre de la recherche biomédicale évoqué plus haut. Aujourd'hui on parle de « participant » à la recherche, ce qui souligne un rôle actif des personnes, et l'évolution se fait vers un véritable partenariat. Que les participants à la recherche en deviennent des partenaires n'est pas neutre au niveau des valeurs sur lesquelles reposent la recherche et ses infrastructures, ni au niveau de leur conception et de leur organisation. Alors que les protocoles de recherche et les projets de cohortes étaient autrefois conçus et organisés exclusivement entre les scientifiques, fussent-ils de diverses disciplines, et leurs institutions, l'implication de la société civile, d'associations de patients ou d'usagers se fait jour dans les cohortes aujourd'hui. Il peut s'agir de consultation de la population concernée, sous différentes formes, avant même la finalisation du protocole, d'entretiens ou de groupes de discussions autour des objectifs et modalités de la recherche, d'implication de membres d'associations dans la rédaction des informations et consentements, pouvant aller jusqu'à la co-construction de protocoles, enfin de la participation de représentants des groupes concernés dans les instances de gouvernance des cohortes. Cette culture de l'implication, de l'engagement du public, est moins développée en France que par exemple au Canada, mais elle s'accroît et la poussée grandissante de l'implication citoyenne en recherche devient réalité. La dimension éthique de cette « poussée sociétale » est profonde et, dans ce domaine, les grandes cohortes sont aussi un vecteur important du renouveau nécessaire et en marche du positionnement de la recherche dans la société.

### Valorisation scientifique des cohortes : une question d'éthique ou une question pratique ?

Les questions qui s'attachent à l'utilisation des cohortes par de nombreuses équipes à travers le partage des données et à la reconnaissance du travail spécifique à cette mise à disposition ont un socle éthique. Mais au-delà des analyses de valeurs, il faut aussi s'intéresser aux obstacles au partage, aux mesures incitatives et au développement d'outils pour rendre opérationnelle la reconnaissance de ce travail dans les infrastructures, notamment les cohortes.

Les cohortes financées souvent sur des fonds publics s'accompagnent de l'impératif de restituer leurs avancées, réalisées par les nombreuses équipes qui les utilisent, auprès de leurs participants et de la société plus

généralement. Des politiques de libre accès aux données et aux résultats doivent être favorisées afin de faciliter la dissémination des informations scientifiques. Cette politique d'incitation au partage se heurte néanmoins à la confidentialité et à la sensibilité de certaines données, nous l'avons vu. Il s'agit donc davantage de trouver un équilibre entre procédures d'utilisation de la ressource raisonnablement contraignantes, destinées à limiter les usages détournés (contrats, accords, chartes...), et mesures incitant à la diffusion et au partage.

L'étude de cohorte peut générer une masse d'informations médicales et scientifiques (bases de données) et d'importantes collections de matériel biologique (parfois structurées en biobanques). Au regard des longues périodes de suivi propres aux cohortes, les données et/ou échantillons sont extrêmement précieux pour la communauté biomédicale et scientifique, et leur partage déterminant pour optimiser la production des connaissances. Or la constitution et l'entretien des bases de données et des collections exigent un effort considérable de la part des chercheurs, techniciens et/ou des professionnels de la santé, qui n'est pas reconnu dans le système d'évaluation actuel. Cet état de fait peut générer des résistances au partage, non seulement à cause de la protection des personnes et de la confidentialité des

données, mais aussi par le sentiment de perte de contrôle, et de non-reconnaissance d'un travail souvent fastidieux et au service de la communauté. Pour attribuer une reconnaissance juste, il est nécessaire d'être capable de mesurer l'apport d'une cohorte. Quels sont les principaux obstacles ?

Un certain nombre d'indicateurs d'activité manquent globalement (encadré ci-dessous) : des indicateurs d'organisation, qui permettraient aux cohortes de prouver leur efficacité et leur utilité ; des indicateurs de qualité, à différents niveaux (échantillons biologiques, données associées...), de caractéristiques des échantillons et des données ; et des indicateurs de production scientifique. Ces derniers sont sans doute les plus appropriés pour démontrer l'impact en recherche de la sollicitation des cohortes. Chacun de ces indicateurs a cependant ses limites et le suivi de l'utilisation d'une cohorte n'est pas évident (encadré page suivante). En particulier, le traçage des publications exploitant une cohorte donnée et le calcul de l'impact en résultant n'est pas facile. On manque notamment d'un procédé de traçage fiable, aisé et automatisé des utilisations des cohortes (et des bioressources en général). Il est nécessaire d'optimiser et de standardiser ce procédé en créant un système d'identifiant unique et traçable pour chaque cohorte et en organisant son utilisation

## Exemples d'indicateurs décrivant un usage et un management efficaces des bioressources générées par les cohortes

### Indicateurs de management

- Nombre de projets financés ou facilités/an
- Quantité de données ou nombre d'échantillons ajoutés/quantité de données ou nombre d'échantillons utilisés pour des projets/an
- Durée moyenne entre collecte et utilisation des données ou de l'échantillon (maintenance durable)

### Indicateurs de qualité

- Participation à des programmes d'évaluation externes tels que certification ou accréditation (ISO, par exemple)
- Détermination de l'étendue et de la richesse des données collectées (par exemple, temps de suivi moyen des sujets)

### Indicateurs de productivité scientifique

- Facteur d'impact cumulatif des publications qui résultent de la recherche soutenue par la cohorte
- Nombre d'exploitations objectivables résultant de l'utilisation de la ressource (bio-marqueurs entrés en pratique, recommandations par des autorités de santé issues des résultats de l'analyse de la cohorte...)
- Nombre de contrats/accords de collaboration signés par an



### Éléments clés nuisant au traçage de l'utilisation des cohortes dans la littérature scientifique

#### Difficultés liées à l'identification et aux remerciements d'une cohorte

- Multiplicité des sections où les cohortes peuvent être remerciées (*Material and methods, Acknowledgements, References...*)
- Remerciements de la cohorte ou citation placées en dehors du titre ou du résumé, dans le corps principal de l'article (ou dans le « Matériel supplémentaire » en ligne), ce qui ne peut être détecté que si on a accès à l'article entier, et n'est pas indexé dans les bases (comme *PubMed* ou *Web of Science*)
- Erreurs de frappe ou approximation du nom de la cohorte
- Multiplicité des noms pour une cohorte donnée
- Utilisation des données en cascade
- Remerciements des personnes à la place de la cohorte
- Absence de remerciements (négligence)
- Pas de méthode standardisée pour inciter les chercheurs à remercier rigoureusement la ressource utilisée dans les recommandations aux auteurs...

#### Difficultés rencontrées avec les publications de référence décrivant les cohortes (« *marker papers* »)

- Système approprié pour se référer à une cohorte donnée, mais pas pour les sous-populations ou jeux de données dérivés de la cohorte

systématique. Une approche opérationnelle s'impose par la création d'outils appropriés.

#### Le Brif, un outil de valorisation

Le Brif (*Bioresource Research Impact Factor*) est un cadre proposé pour répondre à cette problématique [5]. Le premier objectif du Brif est de construire un index quantitatif pour évaluer l'impact d'une bioressource (telle qu'une cohorte) en recherche. Brif s'inspire en partie du *journal impact factor* et fournira une méthodologie et des recommandations pour optimiser la reconnaissance des bioressources, leur utilisation et leur partage à l'échelle internationale. Ce facteur permettrait de documenter : l'utilisation quantitative d'une ressource, le type et l'importance des résultats de recherche l'utilisant, les efforts scientifiques et organisationnels accomplis par ceux et celles ayant mis à disposition par exemple une cohorte de qualité ; par ricochet, ce facteur fournirait également un moyen de reconnaissance aux institutions promotrices de la cohorte. Un tel cadre pourrait être utilisé de manière beaucoup plus formelle qu'une simple « réputation » pour évaluer l'activité d'une cohorte sur le long terme. Une fois pris en compte dans le calcul de la contribution scientifique des chercheurs/managers, le Brif favoriserait la qualité et le partage des ressources par le jeu d'un cercle vertueux : plus grande sera la qualité de la ressource, plus fréquentes seront

les sollicitations ; plus il y aura partage, plus l'impact des contributeurs augmentera, plus il y aura motivation à partager. De plus, la communication de l'identifiant de la cohorte aux participants leur permettrait aussi de pouvoir, s'ils le désirent, tracer eux-mêmes les utilisations de la ressource à laquelle ils participent, leur donnant ainsi une autonomie supplémentaire.

Un groupe de travail international a été créé pour construire le Brif à travers l'étude de quatre axes prioritaires : 1/la création d'un identifiant numérique, 2/l'identification et la hiérarchisation des paramètres, mesures et indicateurs à prendre en compte dans le calcul du Brif, 3/la constitution de recommandations pour citer et référencer ces bioressources de façon standardisée dans les publications scientifiques et 4/l'élaboration ou l'adaptation d'un cadre réglementaire pour l'accès et le partage des ressources (<http://www.gen2phen.org/groups/brif-bio-resource-impact-factor>).

Cet exemple d'un travail en cours montre les liens qui existent entre les dimensions éthiques des cohortes aujourd'hui du point de vue de la protection des personnes, de l'implication des participants et de la société et de la pratique même de la recherche où le partage des données est porteur de valeurs. Les cohortes dans la recherche dans le domaine de la santé sont un domaine privilégié de l'expression de l'ensemble de ces dimensions éthiques. ▮