



HAUTE AUTORITÉ DE SANTÉ

ÉVALUER

LES TECHNOLOGIES DE SANTÉ

GUIDE
MÉTHODOLOGIQUE

Études en vie réelle pour l'évaluation des médicaments et dispositifs médicaux

Validé par le Collège le 10 juin 2021

Descriptif de la publication

Titre	Études en vie réelle pour l'évaluation des médicaments et dispositifs médicaux
Méthode de travail	Guide méthodologique élaboré à partir d'une revue de la littérature, d'un état des lieux des recommandations françaises et internationales et de l'avis d'experts réunis dans un groupe d'appui méthodologique.
Objectif(s)	Ce guide méthodologique vise à soutenir et accompagner les études en vie réelle des produits de santé évaluées par les commissions d'évaluation de la HAS. Il vise à proposer des points de repère pratiques sur les aspects méthodologiques pour optimiser le niveau de preuve de ces études et la confiance dans leurs résultats.
Cibles concernées	Industriels et prestataires
Demandeur	Autosaisine
Promoteur(s)	Haute Autorité de santé (HAS)
Pilotage du projet	Service évaluation des médicaments
Recherche documentaire	Service documentation veille
Auteurs	<p>Équipe projet transversale de la DEMESP : incluant des représentants du service évaluation des médicaments (Judith Fernandez, Julie Taransaud, Mathilde Grande), du service évaluation des dispositifs médicaux (Corine Collignon, Camille Marguerite), du service évaluation médico-économique et de santé publique (Anne-Line Couillerot-Peyrondet, Sophie Tchakamian) et de la mission internationale (Irena Guzina, Chantal Guilhaume). Une chargée de mission externe (Clémentine Nordon) a également été recrutée pour ce projet.</p> <p>Groupe d'appui méthodologique : François Gueyffier, Antoine Vanier, Dominique Costagliola, Sarah Zohar, Pauline Chauvin, Driss Berdaï, Emmanuel Oger, Bruno Falissard, Jean-Luc Faillie, Muriel Vray, Benoit Dervaux, Julia Bonastre.</p>
Conflits d'intérêts	Les membres du groupe de travail ont communiqué leurs déclarations publiques d'intérêts à la HAS. Elles sont consultables sur le site https://dpi.sante.gouv.fr . Elles ont été analysées selon la grille d'analyse du guide des déclarations d'intérêts et de gestion des conflits d'intérêts de la HAS. Les intérêts déclarés par les membres du groupe de travail ont été considérés comme étant compatibles avec leur participation à ce travail.
Validation	Version du 10 juin 2021
Actualisation	

This document and its bibliographic reference are available to download at www.has-sante.fr 

Haute Autorité de santé – Service communication et information
5 avenue du Stade de France – 93218 Saint-Denis La Plaine Cedex. Tél. : +33 (0)1 55 93 70 00
© Haute Autorité de santé – juin 2021 – ISBN : 978-2-11-162649-2

Sommaire

Introduction	4
1. Pourquoi mettre en place une étude en vie réelle ?	7
1.1. Pour décrire les conditions d'utilisation	7
1.2. Pour mesurer l'efficacité et les risques liés à l'utilisation du produit de santé en condition réelle d'utilisation	9
1.3. Pour estimer les scores d'utilité en vie réelle	12
1.4. Pour mesurer la consommation de ressources en vie réelle	13
1.5. Pour mesurer l'impact organisationnel du produit de santé	13
2. Comment réaliser une étude en vie réelle en vue d'une (ré)évaluation par la HAS d'un médicament ou d'un DM ?	15
2.1. Rédiger un protocole, avec le soutien d'un comité scientifique	15
2.2. Proposer un type d'étude en adéquation avec les questions de recherche identifiées	16
2.3. Capitaliser sur les données préexistantes et en particulier celles issues du SNDS	17
2.4. Collecter des données de qualité	17
2.5. Intégrer les mesures d'intérêt pour les patients dans les critères de jugement de l'étude	18
2.6. Garantir la transparence des données	19
3. Références et collaborations internationales	20
3.1. Standards méthodologiques	20
3.2. Collaborations et références internationales	20
Table des annexes	22
Recherche documentaire	43
Références bibliographiques	44
Abréviations et acronymes	50

Introduction

Contexte

La Haute Autorité de santé (HAS) est chargée d'évaluer scientifiquement l'intérêt des médicaments et des dispositifs médicaux (DM)¹ sur le plan médical, économique et de santé publique. Cette évaluation est destinée à éclairer la décision des pouvoirs publics pour la prise en charge des produits de santé concernés par l'assurance maladie, ainsi qu'à déterminer les conditions de leur bon usage, leur place dans la stratégie de prévention, de diagnostic ou thérapeutique, et leur efficacité en vue de la négociation de leur prix.

Les évaluations des produits de santé par la HAS sont réalisées par des commissions spécialisées : la commission de la transparence (CT), la commission nationale d'évaluation des dispositifs médicaux et des technologies de santé (CNEDiMTS), la commission technique des vaccins ainsi que la commission d'évaluation économique et de santé publique (CEESP) dans certains cas. Les principes d'évaluation et méthodes d'analyse scientifiques sont décrits dans des documents spécifiques (1-5).

Lors des demandes d'évaluation en vue du remboursement, ces commissions s'appuient principalement sur les essais cliniques qui sont fondamentaux pour démontrer l'efficacité du produit de santé, conformément aux principes de *evidence based medicine*.

Pour autant, les commissions de la HAS ont une longue histoire d'utilisation des données observationnelles également appelés « en vie réelle », c'est-à-dire les données concernant l'utilisation, l'efficacité ou la tolérance d'un produit de santé en pratique courante. Afin de réaliser sa mission, d'enrichir et d'accompagner l'évaluation des produits de santé, la HAS analyse régulièrement des données observationnelles et sollicite également des données en vie réelle complémentaires aux essais cliniques lorsqu'elles sont indispensables à une réévaluation. En effet, la force de l'essai clinique constitue également sa faiblesse. À titre illustratif, la rigueur du schéma expérimental, qui est nécessaire à la démonstration de l'efficacité, peut entraîner des conditions parfois éloignées de la pratique clinique et remettre en cause la transposabilité des résultats à la pratique courante. Les résultats obtenus sur des durées courtes peuvent être peu compatibles avec l'évolution rapide des technologies et des stratégies.

Alors que l'évaluation historique reposait sur le seul essai clinique, les faits récents montrent que ce modèle évolue dans les évaluations de la HAS. Les nouveaux designs d'essais cliniques dans le contexte du développement de thérapies géniques et de technologies embarquant de l'intelligence artificielle avec des procédés d'apprentissage automatique génèrent des incertitudes sur le déploiement de ces solutions innovantes en pratiques courantes et renforcent les attentes vis-à-vis des études en vie réelle. Le DM devient lui-même un outil de collecte d'information qui peut être utilisé à des fins d'études, dès lors que la réglementation en matière d'utilisation des données de santé est respectée (6-8).

Les promesses sont fortes et elles devront être tenues sur du moyen et long terme. Il s'agit désormais de suivre le maintien ou d'optimiser la performance du produit de santé en vie réelle, d'identifier les patients répondeurs, de comprendre et d'anticiper les toxicités ou risques et d'accompagner l'impact organisationnel.

¹ Dans ce guide méthodologique, le mot DM regroupe les dispositifs médicaux et produits de santé mentionnés aux articles L. 5211-1 ou L. 5221-1 du Code de la santé publique (CSP).

La prise en compte des données collectées en conditions réelles d'utilisation est donc un enjeu majeur dans l'évaluation des produits de santé. En améliorant leur recueil, stockage, analyse et transparence, et plus globalement la confiance que l'on peut porter à leur résultat, la HAS a la conviction que la pertinence de l'évaluation des produits de santé au service des patients ne sera que renforcée.

Il y a un enjeu de méthode car la qualité des études en vie réelle et la validité scientifique de leurs résultats sont des enjeux cruciaux pour leur prise en compte optimisée dans l'évaluation des produits de santé par la HAS (9).

La HAS a donc entrepris la mise à jour du guide méthodologique dédié aux études en vie réelle (10) dans un contexte :

- d'accroissement de l'accessibilité aux données de santé notamment via l'expansion du *Health Data Hub* et d'évolution des méthodes d'analyse augmentant les possibilités d'études comparatives en vie réelle ;
- d'intégration de la perspective des patients et usagers dans l'évaluation des produits de santé qui disposent d'un savoir spécifique sur leur maladie (notamment via la collecte de données directement évaluées par les patients, *Patient Reported Outcome/Experience Measures*) (1, 2, 11) ;
- et de prise en compte de l'importance de disposer de données complémentaires à celles des essais cliniques grâce aux études en vie réelle demandées par la HAS ou réalisées à l'initiative de l'entreprise² du médicament ou du DM (12, 13).

Objectifs du guide

Ce guide utilise le terme de données « en vie réelle » correspondant à l'ensemble des données générées dans des études observationnelles, en conditions de pratique courante.

Ce guide méthodologique vise à soutenir et accompagner la réalisation d'études en vie réelle des produits de santé, en vue de leur évaluation par les commissions d'évaluation de la HAS. Il vise à proposer des points de repère pratiques sur les aspects méthodologiques pour optimiser le niveau de preuve de ces études et la confiance dans leurs résultats.

Ce guide s'applique donc à toutes les études en vie réelle sur les produits de santé (médicaments ou DM), y compris celles demandées par les commissions d'évaluation de la HAS (« études post-inscription »). Il s'adresse à l'ensemble des acteurs impliqués dans la conception et la réalisation des études en vie réelle sur les produits de santé : les industriels, les sociétés de recherche contractuelle, mais aussi les organisations professionnelles et les équipes académiques, qui sont de plus en plus souvent impliquées, notamment dans le cadre de partenariats publics-privés.

Ce guide se focalise sur :

- les études observationnelles ou en vie réelle, qui représentent la majorité des études complémentaires demandées par la HAS ;
- les essais pragmatiques, dans la mesure où ils s'inscrivent dans un continuum entre les essais cliniques randomisés et les études observationnelles.

Ce document n'abordera ni les essais interventionnels non pragmatiques (essais cliniques classiques), ni les méta-analyses ou les méta-analyses en réseau, bien qu'ils puissent être demandés dans le cadre d'études post-inscription. Par ailleurs, les aspects méthodologiques des données en vie réelle sur l'épidémiologie, le fardeau de la maladie ou celles constituées à des fins de comparaisons indirectes ne

² Dans ce guide, par simplification, le terme entreprise peut également désigner les fabricants et les distributeurs.

seront pas détaillés dans ce guide bien que ce type de données puissent également contribuer aux évaluations de la HAS.

Le guide est articulé en trois chapitres :

- le premier aborde les questions de recherche pouvant survenir lors d'un développement clinique d'un médicament ou d'un DM et qui peuvent justifier la mise en place d'une étude en vie réelle ;
- le deuxième détaille les principales recommandations de la HAS pour réaliser une étude en vie réelle de qualité ;
- et le troisième précise les références méthodologiques internationales à prendre en compte pour la réalisation d'une étude en vie réelle.

Pour les modalités pratiques concernant les échanges de document avec la HAS et les procédures internes relatives aux études post-inscription, des notices dédiées sont disponibles sur le site internet de la HAS :

- [pour les DM](#) ;
- [pour le médicament](#) ;
- pour les produits de santé faisant également l'objet de demandes de données complémentaires de la part de la CEESP, le protocole est amené à être discuté avec le service évaluation économique et santé publique (SEESP) de la HAS. Le service peut aussi être contacté pour des dépôts de protocoles spécifiques à l'analyse économique ou étude organisationnelle.

Ce document ne donne pas de « recette toute prête », pouvant être appliquée quelle que soit la nature de la demande et doit être considéré comme une aide méthodologique. Il revient à l'industriel de proposer et d'adapter son protocole pour mettre en place une étude en vie réelle pertinente en vue d'une évaluation des produits de santé par la HAS. Il n'est pas un document opposable.

1. Pourquoi mettre en place une étude en vie réelle ?

Cette section détaille les questions de recherche pouvant survenir lors d'un développement clinique d'un médicament ou d'un DM et qui peuvent justifier la mise en place d'une étude observationnelle en vue d'une évaluation du produit de santé par la HAS.

Ce paragraphe concerne principalement les données générées après la mise sur le marché en France mais il est toutefois rappelé que les données en vie réelle sont également prises en compte dans les premières évaluations des produits de santé par la HAS, notamment celles générées dans le cadre des accès précoces ou des prises en charge transitoires. De plus, il est rappelé que des données en vie réelle sur l'épidémiologie, le fardeau de la maladie ou celles constituées à des fins de comparaisons indirectes peuvent également contribuer aux évaluations de la HAS.



1.1. Pour décrire les conditions d'utilisation



- ➔ À qui et comment le produit de santé est-il prescrit en conditions réelles d'utilisation ?
- ➔ Y a-t-il une prescription préférentielle du produit de santé à des patients ayant des caractéristiques particulières par rapport à ceux qui reçoivent les comparateurs ?

Population cible et population rejointe

La différence potentielle entre la population cible (éligible au remboursement selon l'évaluation de la HAS) et la population rejointe (recevant réellement ce produit de santé) est l'un des éléments d'incertitude dans l'évaluation des produits de santé. Il existe souvent une différence entre les caractéristiques des participants inclus dans les essais cliniques classiques – qui eux-mêmes sont souvent un sous-

ensemble de la population cible – et celles des patients qui reçoivent ce produit de santé en pratique courante de soins, c'est-à-dire la population rejointe. En effet :

- les centres d'investigation internationaux participant aux essais cliniques classiques ne sont pas toujours représentatifs des centres français prenant en charge les patients en pratique courante ; les patients qu'ils prennent en charge sont également différents ;
- les critères d'éligibilité des participants aux essais cliniques classiques sont plus stricts et restrictifs que ceux de l'AMM ou de la population éligible au remboursement (14-16) ou du marquage CE, ce qui questionne la représentativité de la population sur laquelle repose l'estimation de l'efficacité du produit de santé par rapport à la population cible française. Par exemple, les patients ayant un âge avancé, des comorbidités, prenant d'autres médicaments, ayant une grossesse en cours, etc., sont très souvent non éligibles dans les essais cliniques et retenus dans le périmètre de remboursement ;
- les caractéristiques socio-démographiques des patients des essais cliniques, en particulier le ratio hommes/femmes des essais cliniques, l'âge ou encore l'ethnicité des patients, ne reflètent pas forcément celles observées chez les patients recevant ou utilisant effectivement le produit de santé. Par conséquent, ces données ne permettent pas toujours d'analyser les potentielles différences liées aux caractéristiques socio-démographiques et/ou d'investiguer l'efficacité et la tolérance d'un produit de santé dans le groupe moins représenté (17) ;
- les patients recevant réellement le traitement peuvent le recevoir en dehors du périmètre remboursable (18-21).

Ainsi, il peut être utile de caractériser la population rejointe en identifiant quels patients sont réellement traités en pratique courante, que ce soit dans le cadre du périmètre de remboursement ou non. Une description du parcours de soin, et notamment des centres prenant en charge les patients d'intérêt, peut être également attendue.

Par ailleurs, en pratique courante, les médecins ne prescrivent pas tel ou tel produit de santé de façon aléatoire, comme dans un essai clinique. Ils peuvent préférentiellement prescrire le nouveau produit de santé à certains patients plutôt qu'à d'autres, en fonction des antécédents des patients, de leurs traitements passés, actuels ou de facteurs pronostiques (liés, par exemple, à l'âge, à la sévérité de la maladie, à la fréquence des effets secondaires, etc.). Cette prescription préférentielle, qui peut également être due aux préférences des patients ou des aidants, peut être investiguée dans une étude en pratique courante afin de mieux caractériser la population rejointe. Par ailleurs, cette propension des patients à recevoir tel ou tel traitement doit être prise en compte lors de la recherche d'un lien causal entre produit de santé et effet du produit mesuré pour minimiser les biais (biais d'indication ou biais de sélection).

Modalités d'utilisation

La différence entre les modalités d'utilisation du produit de santé par les patients ou de sa prescription par les médecins, entre les essais cliniques classiques d'une part et la pratique courante de soins d'autre part, constitue un autre niveau d'incertitude.

Lorsqu'un produit de santé est utilisé par le patient lui-même, il peut exister une différence entre les modalités d'utilisation du produit de santé (niveau d'observance et de persistance, interruptions) dans les essais cliniques classiques où la prise du produit de santé est encadrée, particulièrement suivie et pendant un temps généralement limité, et l'utilisation en pratique courante de soins. Les patients ayant un faible niveau d'observance thérapeutique sont souvent non éligibles aux essais cliniques classiques. Par ailleurs, pendant le déroulé de l'essai clinique, des mesures de surveillance et d'amélioration de

l'observance thérapeutique peuvent être mises en place (prises de sang, rappels faits au patient, visites programmées). Au contraire, en pratique courante de soins, les patients auront un niveau d'observance thérapeutique habituel. Une mauvaise observance thérapeutique peut être liée au patient lui-même et/ou au produit de santé (pénibilité, pas d'efficacité ressentie ou au contraire sentiment d'aller mieux et d'être guéri, effets indésirables, etc.). Lorsque le produit de santé est utilisé par un professionnel de santé, il peut exister une différence dans les modalités d'utilisation entre les professionnels de la pratique courante et ceux de l'essai clinique, mieux expérimentés. L'impact de l'environnement, l'expertise des centres utilisateurs, le plateau technique ou l'organisation de la prise en charge sont ainsi des éléments à prendre en compte lors de l'analyse de la transposabilité des résultats d'une étude clinique à la pratique courante.

Concernant les prescripteurs, il peut exister une différence entre les modalités de prescription du produit de santé dans les essais cliniques classiques et ces modalités en pratique courante de soins. Cette différence peut concerner le dosage, le nombre de prises quotidiennes, la durée recommandée de traitement, les arrêts de traitement, la prise en compte des interactions avec d'autres médicaments, des contre-indications et des recommandations de surveillance du traitement. À noter que cette différence peut refléter la nécessité pour les prescripteurs d'ajuster les doses afin d'optimiser le ratio efficacité/tolérance du traitement – nécessité qui n'avait pas pu être mise en évidence lors des essais cliniques.

Pour répondre à ce questionnement, il convient de décrire les modalités de prescription et d'utilisation des produits de santé en pratique courante.

1.2. Pour mesurer l'efficacité et les risques liés à l'utilisation du produit de santé en condition réelle d'utilisation



- ➔ Est-ce que l'efficacité évaluée dans les essais cliniques sera la même en condition réelle d'utilisation ? Est-ce que la quantité d'effet mise en évidence dans les essais cliniques sera observée en pratique courante ?
- ➔ Est-ce que les bénéfices et risques tels que mis en évidence dans les essais cliniques seront transposables à d'autres contextes : centres moins expérimentés, patients et système de soins français ?
- ➔ Est-ce que l'impact sur la qualité de vie ou toute autre mesure pertinente pour les patients mise en évidence dans les essais cliniques seront les mêmes en conditions réelles d'utilisation ?

Du fait des différences de population, de durée de suivi, de modalités d'usage et de prescription du produit de santé (cf. *supra*), il peut exister une différence entre l'effet du produit de santé tel que mesuré lors des essais cliniques classiques et celui effectivement observé en vie réelle. Cette différence peut porter :

- sur l'effet bénéfique attendu en termes de morbi-mortalité (*efficacy-effectiveness gap*) (22) ;
- sur l'impact du produit de santé sur la qualité de vie liée à la maladie ou toute autre mesure pertinente pour les patients ;
- ou sur les risques et la sécurité d'emploi du produit de santé.

Cette différence provient essentiellement du fait que certains éléments de contexte des soins ou d'utilisation en vie réelle – et non reflétés dans les essais cliniques – peuvent modifier l'effet propre du produit. Ces éléments de contexte agissant comme des modificateurs de l'effet propre du produit de santé peuvent être liés au système de santé (par exemple, l'accessibilité des patients aux soins), aux modalités de prescription et d'utilisation (par exemple, doses effectivement utilisées ou interactions médicamenteuses, caractère opérateur dépendant pour le DM, organisation des soins) et/ou liés aux caractéristiques du patient en pratique courante (par exemple, patients plus hétérogènes que dans les essais cliniques en termes de comorbidités, d'âge ou d'observance thérapeutique) (23, 24).

Ces incertitudes peuvent remettre en question la validité externe des essais cliniques (25, 26) et nécessiter la mise en place d'une collecte de données observationnelles en conditions réelles d'utilisation.

2

- **L'efficacité du produit de santé, par rapport aux comparateurs pertinents dans la stratégie en France, sera-t-elle de même importance que celle mise en évidence dans les essais cliniques ?**
- **L'efficacité, la sécurité d'emploi du produit et/ou toute autre mesure pertinente pour les patients comparativement à la stratégie de référence confirmeront-elles l'intérêt évalué sur la base d'essais cliniques non comparatifs ?**
- **Quels seront les effets du produit de santé par rapport à des comparateurs pertinents non pris en compte lors de l'évaluation initiale, notamment ceux issus des développements cliniques concomitants ?**

L'accélération des développements cliniques provoquant la mise sur le marché de produits de santé de façon précoce, sur la base d'essais cliniques non comparatifs ou de développement d'essais concomitants, renforce la nécessité d'analyser l'efficacité et la sécurité relative d'un produit pour s'assurer de son intérêt thérapeutique en conditions réelles d'utilisation. La question posée est alors celle de l'efficacité comparative du produit de santé d'intérêt.

La question de l'effet du produit de santé par rapport à l'ensemble des comparateurs pertinents se pose également lorsqu'il s'agit d'évaluer l'efficacité du produit de santé ; l'ensemble des comparateurs pertinents pouvant ne pas être inclus dans l'analyse économique au moment de l'évaluation initiale, notamment en raison des données disponibles. En effet, le comparateur du produit de santé d'intérêt dans un essai clinique peut être un placebo, ou un traitement actif mais peu pertinent dans le contexte de soins français au moment de l'évaluation. Par ailleurs, des comparaisons indirectes sont parfois déposées dans les dossiers de demande de remboursement afin de prendre en compte les comparateurs pertinents ; cependant, les comparateurs retenus dans les comparaisons indirectes peuvent ne pas être exhaustifs par rapport à la stratégie thérapeutique en France.

L'objectif d'une étude en vie réelle peut être de mesurer l'efficacité comparative du produit de santé par rapport aux comparateurs pertinents dans le contexte français. Par ailleurs, des données en vie réelle peuvent être générées ou rassemblées pour servir de groupe contrôle externe pour un essai clinique non randomisé. Dans ce cas, la mise en place d'une comparaison externe devra être anticipée et prévue *a priori* pour en améliorer la robustesse et l'inscrire dans une démarche hypothético-déductive (27).

2

→ Quels seront les effets et risques du produit de santé sur le long terme ?

À l'issue d'un essai clinique, les effets et risques à long terme ou de survenue rare ne sont pas ou peu connus. En effet, les effets du produit de santé sur la survie globale, survie sans progression ou sur la survenue de tout type d'évènement sont estimés sur des durées d'essai en général non suffisantes pour valider les hypothèses d'extrapolation dans les modèles économiques, notamment sur les hypothèses de maintien de l'effet traitement ou de longévité d'un implant. L'objectif d'une étude en vie réelle peut être d'estimer les effets à long terme du traitement afin de valider les hypothèses d'extrapolation faites au moment de la mise sur le marché du produit de santé ou de quantifier les reprises chirurgicales, les complications. Il s'agira alors de quantifier l'impact de ces hypothèses sur les résultats d'efficacité, de sécurité et d'efficience. La collecte de données d'efficacité et/ou de tolérance à long terme peut également être demandée en vue d'une réévaluation clinique d'un produit de santé, notamment quand il existe des incertitudes majeures sur le devenir des patients/utilisateurs.

2

→ Quelle est la pertinence de l'effet du produit pour les patients ?

Dans les essais cliniques classiques, le critère de jugement principal doit être objectif, reproductible, bien défini et mesuré de la même façon par tous les investigateurs de l'essai, afin d'éviter les biais de classification. Ce critère est le plus souvent basé sur un élément clinique objectivable (par exemple, une hospitalisation, un évènement clinique tel qu'un infarctus du myocarde, etc.), voire sur un biomarqueur directement lié au mécanisme d'action de celui-ci (par exemple, dosage d'un marqueur tumoral, taux de cholestérol) (28). Or pour les patients, l'intérêt d'un produit de santé ne réside pas uniquement dans l'amélioration clinique ou biologique de la maladie, d'autant plus que ces aspects ne sont pas forcément les plus pertinents de leur point de vue. L'incertitude sur le produit de santé liée à sa pertinence du point de vue des patients peut donc justifier la mise en place d'une étude en vie réelle. Cette incertitude va au-delà du simple ressenti des symptômes (par exemple, douleur, fatigue). Ces attentes et les conséquences de la maladie sur leur vie peuvent être relatives à l'organisation de leur vie quotidienne, la peur de l'avenir, la pénibilité du traitement et bien d'autres aspects.

Les dimensions pertinentes pour les patients ont longtemps été regroupées sous le terme générique de « qualité de vie liée à la santé », qui prend en considération de nombreux aspects relatifs à la perception d'une personne de la qualité de vie affectée par sa maladie ou son handicap : aspects physiques, psychologiques et sociaux. Néanmoins, d'autres mesures peuvent être pertinentes pour les patients comme la satisfaction vis-à-vis des soins, l'impotence fonctionnelle ou l'observance thérapeutique.

1.3. Pour estimer les scores d'utilité en vie réelle

3

→ Quelle est la qualité de vie, valorisée par des scores d'utilité, des patients traités en conditions réelles d'utilisation ?

Un autre niveau d'incertitude concerne la valorisation des résultats sur un état de santé atteint grâce au produit de santé et décrit par l'EQ-5D, lorsque ces résultats sont valorisés par des scores d'utilité³. Le constat d'un défaut de documentation de ce critère peut être fait au moment de l'évaluation initiale.

En effet, au moment de la mise sur le marché, la mesure et la valorisation des états de santé des patients peuvent ne pas être robustes ou ne pas respecter les recommandations en vigueur (3) en particulier lorsque l'essai clinique pivot ne permet pas d'obtenir des données de qualité de vie permettant de valoriser les états de santé. C'est le cas, par exemple, lorsque :

- l'essai n'a pas prévu un recueil de données de qualité de vie liée à la santé avec le questionnaire recommandé (ex. : l'EQ-5D-5L dans la version actuelle du guide HAS de l'évaluation économique) (3) ;
- la durée de suivi de l'essai pivot, son effectif ou les données manquantes ne sont pas compatibles avec la mesure du score d'utilité associée à certains états de santé ;
- l'essai clinique aboutit à des résultats inattendus ou sans effets significatifs sur la qualité de vie en raison d'un effectif trop faible (maladie rare) ou de l'impossibilité de recueil au moment où survient un événement d'intérêt (par exemple, crise aiguë ou événement ponctuel).

Dans ces situations, l'estimation des scores d'utilité peut être non robuste ou biaisée. La revue de la littérature conduite pour pallier l'absence de données issues d'un essai ne permet pas toujours d'identifier une source adaptée à la situation et peut conduire, par exemple dans le cadre de l'évaluation initiale, à utiliser des scores d'utilité valorisés avec une matrice de pondération étrangère ou des scores d'utilité mesurés sur une autre population que celle étudiée ou, encore, se fonder sur des hypothèses fortes quant à l'équivalence entre différents états de santé. Enfin, des événements intercurrents, tels que les événements indésirables d'intérêt, ont pu ne pas être pris en compte en termes de pertes de qualité de vie (désutilité).

Une étude en vie réelle peut avoir pour objectif d'estimer de façon robuste le score d'utilité d'un état de santé en particulier, comme un état atteint à plus long terme, ou la désutilité associée à un événement d'intérêt. En pratique courante, il s'agit notamment de recueillir la perception de la qualité de vie associée à des états de santé caractéristiques de la maladie par le questionnaire EQ-5D-5L auprès d'un échantillon représentatif de la population réellement traitée en France afin de pouvoir les valoriser par la matrice de pondérations de la population française (3).

³ Dans le cadre des évaluations économiques, lorsque la qualité de vie liée à la santé est identifiée comme une conséquence importante du produit de santé évalué, la durée de vie pondérée par la qualité de vie liée à la santé est le critère de résultat de santé retenu ; la qualité de vie liée à la santé étant mesurée par un score d'utilité (cf. guide).

1.4. Pour mesurer la consommation de ressources en vie réelle

4

- La mesure des ressources consommées en pratique courante est-elle différente de celle considérée au moment de la mise sur le marché du produit de santé ? Y a-t-il des ressources consommées qui n'avaient pas été identifiées ou, au contraire, des ressources consommées lors de l'essai pivot mais non observées en conditions réelles d'utilisation ?
- Si la population traitée en vie réelle est différente de la population cible considérée au moment de la mise sur le marché du produit de santé ou si les pratiques ou le contexte de soins évoluent, quel est l'impact de cet écart sur la mesure et la valorisation des ressources consommées ?

Au moment de la mise sur le marché, le recueil des ressources consommées peut ne pas couvrir l'éventail complet des ressources qui vont être consommées en conditions réelles d'utilisation (3). En effet, si le recueil a été réalisé au cours d'un essai clinique, les postes de coûts recueillis peuvent être contraints par le protocole, en termes de champ d'analyse (par exemple, coûts hospitaliers uniquement), de temporalité (contraintes de planning, nombre de visites planifié), de durée de suivi (absence de prise en compte du coût associé à l'évolution de l'état de santé ou du handicap si la durée de suivi est courte) ou de perspective (par exemple, coût pour l'hôpital uniquement, et non pour le patient ou pour la collectivité). Le recueil des ressources consommées dans un essai clinique peut aussi avoir été réalisé dans plusieurs pays, ce qui questionne la transposabilité de cette mesure au contexte français. Les données peuvent être aussi de mauvaise qualité car non renseignées.

La mesure des ressources consommées peut également provenir d'études en vie réelle françaises, dont le recueil peut ne pas être complet (absence de recueil de certains postes de coûts, données manquantes) ou qui porte sur une population d'analyse plus ou moins différente de celle qui sera traitée en vie réelle (population rejointe).

1.5. Pour mesurer l'impact organisationnel du produit de santé

5

- Les impacts anticipés au moment de la mise sur le marché d'un produit de santé s'avèrent-ils réels et mesurables par des critères en conditions réelles ?
- Les conditions réelles d'utilisation amènent-elles à révéler des impacts organisationnels non identifiés au moment de la mise sur le marché du produit de santé ?

Au moment de la mise sur le marché d'un produit de santé, des impacts organisationnels sont fréquemment revendiqués : impacts sur l'organisation des soins, les pratiques professionnelles ou les conditions de prise en charge des malades. Ils demeurent cependant rarement documentés ou sont souvent identifiés de façon incomplète au moment de l'évaluation initiale et seulement rapportés de manière descriptive et non argumentés.

L'absence de cadre structuré de définition des effets du produit de santé sur une organisation de santé y participe probablement pour une bonne part. Mieux prendre en compte cette dimension de

l'évaluation des produits de santé nécessite tout d'abord d'en définir les contours. Forte de ce constat, la HAS a élaboré un guide méthodologique (5) qui propose une cartographie dont l'objectif est de structurer la manière dont les impacts organisationnels d'un produit de santé peuvent être identifiés et argumentés. Cette cartographie propose à cet effet une classification composée de macro-critères et de critères assortis d'exemples d'indicateurs.

Une étude en vie réelle peut donc avoir pour objectif d'identifier ou de mesurer de façon objective les impacts organisationnels associés à la mise sur le marché du produit de santé, notamment au travers de la cartographie élaborée par la HAS.

2. Comment réaliser une étude en vie réelle en vue d'une (ré)évaluation par la HAS d'un médicament ou d'un DM ?



2.1. Rédiger un protocole, avec le soutien d'un comité scientifique

La première étape est la rédaction d'un protocole d'étude devant détailler :

- la justification et l'objectif de l'étude ;
- la ou les hypothèses testées pour les études comparatives (notamment pour le calcul du nombre de sujets nécessaires) ;
- les populations sources et analysées ;
- les sources de données ;
- la méthode de l'étude (définition et mesure de l'exposition et des événements d'intérêt, méthodes de minimisation et de contrôle des biais, plan d'analyse statistique) ;
- la conduite de l'étude et le contrôle de la qualité ;
- les limites attendues de l'étude ;
- les aspects éthiques.

Les éventuels amendements au protocole survenant au cours de la réalisation de l'étude doivent être déclarés, dûment justifiés et répertoriés.

Pour la rédaction et la validation du protocole, il est préconisé de mettre en place un comité scientifique qualifié et multidisciplinaire, comprenant également des experts patients ou des représentants d'association d'utilisateurs.

Par ailleurs, un comité d'experts pour l'évaluation des événements d'intérêt (critères d'efficacité et/ou événements indésirables) peut être mis en place.

Pour plus d'information concernant l'élaboration d'un protocole d'une étude en vie réelle, se référer à [l'annexe 1](#) relative aux références internationales.

2.2. Proposer un type d'étude en adéquation avec les questions de recherche identifiées

Le type d'étude doit être déterminé en fonction des questions de recherche identifiées (cf. partie 1 du guide).

Pour répondre à la demande de données complémentaires formulée par la HAS (« étude post-inscription »), il est précisé que la mise en place de plusieurs études peut être utile ou nécessaire. Il appartient à l'entreprise ou au protomoteur de l'étude de justifier le type d'étude retenu et sa capacité à répondre de façon adéquate à la demande de données formulée par la HAS. Enfin, les limites méthodologiques et éventuels biais attendus devront être discutés.

→ Une étude en vie réelle pour documenter l'utilisation d'un produit de santé en pratique courante

Les études observationnelles descriptives et non comparatives sont alors le type d'étude à privilégier. Elles permettent de décrire la population rejointe, les modalités de prescription et d'utilisation d'un produit de santé en conditions réelles d'utilisation (caractéristiques des patients et des prescripteurs, stratégies thérapeutiques, modalités d'instauration ou d'arrêt des traitements, potentiel mésusage, etc.), l'évolution clinique des patients exposés au cours du temps, la qualité de vie ressentie par le patient ou le profil de sécurité d'emploi du produit de santé. En accord avec son rapport d'analyse prospective 2020, la HAS préconise de prendre en compte le sexe dans la construction du projet de recherche et de veiller à ce que le ratio femmes/hommes des patients de l'étude reflète celui de la pratique courante (17). Par ailleurs, ces études apportent une multitude d'informations sur les éléments contextuels qui interagissent potentiellement avec l'effet du produit de santé et le devenir des patients mais ne permettent pas d'affirmer une association causale.

Pour plus d'information, se référer à [l'annexe 2](#) relative aux études en vie réelle descriptives et leurs considérations méthodologiques.

→ Une étude en vie réelle pour analyser l'efficacité, la sécurité et/ou l'efficience d'un produit de santé en comparaison à une autre stratégie thérapeutique

Les études comparatives seront alors le design à privilégier. En plus de pouvoir caractériser la population qui utilise réellement le produit de santé, ces études peuvent en effet estimer l'efficacité, la sécurité ou toute autre mesure pertinente pour les patients en conditions réelles d'utilisation par rapport à leurs comparateurs cliniquement pertinents. Ces études peuvent être des études en vie réelle comparatives, des essais pragmatiques ou des essais cliniques classiques. En l'absence de randomisation, la HAS préconise de mettre en place des mesures de minimisation du biais d'indication, notamment via l'utilisation de méthodes d'inférence causale en situation observationnelle, comme l'ajustement, l'appariement ou la pondération lorsque les conditions d'application de ces méthodes sont remplies. Dans certains cas, des études de comparaisons avant/après peuvent être pertinentes pour décrire l'impact d'un changement de prise en charge.

Pour plus d'information, se référer à [l'annexe 3](#) relative aux études comparatives et leurs considérations méthodologiques. Cette annexe aborde également la modélisation prédictive dont l'objectif est d'extrapoler des observations du présent à un autre contexte.

D'autres types d'études sont possibles en fonction des questions de recherche ou des demandes de données formulées par la HAS ou d'autres agences.

2.3. Capitaliser sur les données préexistantes et en particulier celles issues du SNDS

L'utilisation de données issues de cohortes et registres peut se faire dans le cadre d'une étude *ad hoc* (« données primaires »), ou à partir de bases de données déjà existantes (« données secondaires ») françaises ou internationales.

Bien que ces deux approches soient acceptables, la HAS recommande de capitaliser sur les données déjà existantes dans l'objectif d'accélérer la disponibilité des résultats, de limiter la duplication des collectes de données et donc *in fine* d'améliorer la collecte. La France dispose de données médico-administratives couvrant le territoire national et de cohortes dans différentes maladies de bonne qualité méthodologique, en particulier dans les maladies rares. L'accès à ces données a été facilité ces dernières années et des dispositifs nationaux vont à l'avenir favoriser la relation entre les producteurs de données et les utilisateurs (*Health Data Hub*, France Cohortes, etc.). Des initiatives pour faciliter l'identification des bases de données existantes sont en cours comme le [programme de l'Alliance pour la recherche et l'innovation des industries de santé \(ARIIS\)](#) et du [Comité stratégique de filière industries et technologies de santé](#).

La HAS rappelle par ailleurs l'importance d'anticiper les besoins de données en vie réelle pour l'évaluation des produits de santé par la mise en place de registres maladies, en particulier dans les maladies rares. Ces registres peuvent être réalisés dans le cadre de partenariat public/privé et s'appuyer sur les filières maladies rares.

Enfin, il est précisé que lorsque la HAS demande des données complémentaires (« étude post-inscription »), cela n'implique pas nécessairement la mise en place d'une étude *de novo*. En effet, une étude observationnelle ou interventionnelle déjà mise en place à l'initiative de l'entreprise, d'une autre agence ou d'un conseil national professionnel (ex. : étude PASS ou PAES pour le médicament) peut être susceptible de répondre à la demande de la HAS. Lorsque des données complémentaires demandées par la HAS nécessitent une collecte *ad hoc* de données, la HAS préconise alors de concevoir cette collecte pour permettre un possible chaînage avec le SNDS en vue d'une réutilisation de ces données de santé par d'autres acteurs. Le stockage des bases de données générées par les études post-inscription sur le *Health Data Hub* est également conseillé par la HAS.

Pour plus d'information, se référer à l'[annexe 4](#) relative aux sources de données possibles et leurs considérations méthodologiques. Se référer à l'[annexe 5](#) pour un focus particulier sur les données du SNDS avec une section spécifique concernant les algorithmes pour l'identification des pathologies et les chaînages possibles avec d'autres sources de données.

2.4. Collecter des données de qualité

La production de données de bonne qualité est un des prérequis indispensables pour que les résultats de l'étude soient pris en compte par les commissions d'évaluation de la HAS. En accord avec les standards méthodologiques internationaux, il est notamment recommandé de :

- documenter au mieux la représentativité des centres, des investigateurs et des patients inclus dans l'étude. En conformité avec les exigences règlementaires, la mise en place d'un registre de non-inclusion, lorsqu'elle est possible, ou le croisement des inclusions avec les données d'activité ou de vente sont conseillés pour justifier la représentativité de l'étude ;
- limiter au maximum les patients perdus de vue et les données manquantes à l'aide d'un monitoring et un contrôle qualité adéquat des données. Pour les études *ad hoc*, des vérifications

sur site doivent être prévues afin de contrôler les informations recueillies, sur l'ensemble des dossiers ou sur un échantillon tiré au sort ;

- sauf demande contraire, s'assurer que l'étude permettra de documenter la population rejointe représentative de tous les patients traités par le produit de santé en vie réelle ;
- lors de l'analyse statistique, comparer, pour étayer la représentativité, les caractéristiques des centres investigateurs participants aux non participants (les comparer également à l'échelon national), celles des patients inclus aux patients non inclus dans l'étude et comparer les caractéristiques des patients selon leur statut (perdus de vue ou non, données manquantes ou non concernant les principales variables d'intérêt).

Pour plus d'information, se référer à [l'annexe 1](#) relative aux références méthodologiques internationales.

2.5. Intégrer les mesures d'intérêt pour les patients dans les critères de jugement de l'étude

Pour les patients, l'intérêt d'un produit de santé ne réside pas uniquement dans l'amélioration clinique ou biologique de la maladie. En effet, ces aspects ne sont pas forcément les plus pertinents du point de vue des patients. Par conséquent, la HAS encourage le recours aux critères de jugement à partir de données directement collectées par les patients, que ce soit comme critère de jugement secondaire ou principal, documentant notamment la qualité de vie. Pour les études post-inscription, avec une demande de données de qualité de vie par la HAS, il est préconisé de :

1. intégrer un auto-questionnaire (*Patients Reported Outcome Measures* – PROMs) collectant des données permettant d'analyser la qualité de vie des patients et/ou toutes autres mesures pertinentes pour les patients. La validité et l'interprétation de ce questionnaire dans la maladie investiguée devront être justifiées par une revue de la littérature. Des hypothèses sur les domaines de qualité de vie susceptibles d'être impactés en priorité devront être formulées *a priori* ;
2. prévoir l'utilisation d'une question de type *Patient Global Impression of Change* et/ou d'un questionnaire de qualité de vie générique en l'absence d'auto-questionnaire spécifique validé dans la maladie ;
3. conformément au guide HAS de l'évaluation économique (3) et dans le cadre spécifique d'une analyse de type coût-utilité, privilégier le recours au questionnaire EQ-5D-5L⁴⁵.
4. L'absence d'utilisation d'un PROMs devra être systématiquement justifiée lors du dépôt du protocole de l'étude post-inscription.

Pour plus d'information, se référer à [l'annexe 6](#) relative aux mesures d'intérêt pour les patients (*Patients Reported Outcome Measures*).

⁴ Pour l'évaluation clinique, il faut cependant noter que le questionnaire EQ-5D n'est pas à considérer comme un instrument mesurant les multiples aspects de la qualité de vie.

⁵ Pour les populations pédiatriques, l'utilisation d'un système de mesure générique développé et validé chez les enfants et adolescents est recommandée (ex. : instrument HUI).

2.6. Garantir la transparence des données

Le principe de transparence est de rendre systématiquement publiques toutes les informations sur la recherche, tout en respectant la protection des données personnelles. Il permet de renforcer la confiance des acteurs dans ces données, en mettant à disposition des informations sur le protocole et les éventuels amendements effectués, sur les sources de données et méthodes d'analyse effectuées (décrites suffisamment précisément pour permettre leur réplique), ainsi que sur les résultats de l'étude.

Il est recommandé d'enregistrer l'étude dans une ou plusieurs bases de données publiques, telles que le portail Épidémiologie-France (<https://epidemiologie-france.aviesan.fr/>), le registre européen EUPAS de l'ENCePP (<http://www.encepp.eu/encepp/studiesDatabase.jsp>), le registre américain Clinical-Trial.gov (<https://clinicaltrials.gov/>), le portail de l'Organisation mondiale de la santé (<https://apps.who.int/trialsearch/Default.aspx>) ou encore le portail PROSPERO pour les revues de la littérature et les méta-analyses (<https://www.crd.york.ac.uk/prospéro/>).

Comme rappelé par le code de conduite de l'ENCePP (29), le protocole doit être finalisé et chaque fois que possible rendu public avant que la recherche ne débute. L'enjeu est de pouvoir fournir la garantie que les résultats de l'étude ont bien été produits selon le protocole et le plan d'analyses statistiques prévus initialement, et non orientés par des analyses réalisées en cours d'étude. De plus, la publication scientifique du protocole est également possible et même encouragée par certains éditeurs de revues scientifiques.

La HAS recommande également la publication des résultats dans une revue scientifique et incite les industriels du médicament et du DM à se rapprocher du *Health Data Hub* pour stocker les bases de données des études en vie réelle en vue de leur réutilisation.

Pour plus d'information concernant la publication des informations et données d'une étude, se référer à [l'annexe 1](#) relative aux références internationales.

3. Références et collaborations internationales

3.1. Standards méthodologiques

Ce guide méthodologique a été rédigé en accord avec les références méthodologiques internationales pour la réalisation des études en vie réelle. Le respect des standards méthodologiques et des bonnes pratiques pour la conduite des études en vie réelle est en effet un prérequis indispensable pour que les résultats de l'étude soient pris en compte par les commissions d'évaluation de la HAS.

Afin de réaliser une étude en vie réelle de qualité, il est donc nécessaire de consulter les bonnes pratiques de mise en œuvre des études en pratique courante, depuis la définition de l'objectif jusqu'à la publication des résultats. Les principaux guides de référence sur le développement du protocole, la conduite de l'étude (incluant la collecte des données, le data management, le contrôle de qualité et les aspects éthiques) et sur la rédaction des rapports de résultats sont référencés en annexe 1 de ce document.

3.2. Collaborations et références internationales

Consciente de l'essor des études en vie réelle en France comme à l'international, la HAS est connectée aux autres agences compétentes dans ce domaine, notamment grâce au réseau européen des agences d'évaluation des technologies de santé (*European Network for Health Technologies Assessment – EUnetHTA*). Ce réseau, créé en 2006, a pour objectif d'améliorer les pratiques d'évaluation des technologies de santé et d'instaurer une collaboration efficace et pérenne entre les différents membres du réseau. Les travaux menés au sein du réseau l'ont été sur une base volontaire et dans le cadre d'actions conjointes, bénéficiant d'un soutien financier de la Commission européenne. La troisième action conjointe est en cours depuis 2016, pour une durée de 5 ans.

Ces actions de coopération se déclinent en plusieurs groupes de travail (Work Package/WP), dont un axe spécifique à l'amélioration de la qualité des données produites au cours du développement et du cycle de vie des technologies de santé (WP5). Cet axe, qui porte d'un côté sur l'activité de dialogues précoces (*early dialogues*) et de l'autre, sur les études post-autorisation et le recueil de données complémentaires (*post launch evidence generation*), est coordonné par la HAS dans le cadre de l'action conjointe en cours.

Lors des actions conjointes précédentes, plusieurs outils et documents méthodologiques ont été développés :

- des critères de sélection des technologies pour lesquelles une demande d'étude complémentaire pourrait être justifiée (30) ;
- une base de données (base EVIDENT), permettant de répertorier les demandes d'études complémentaires faites par les agences HTA européennes (accessible uniquement aux agences partenaires du réseau (<https://eunetha.eu/evident-database/>) ;
- deux documents méthodologiques (*position papers*) portant sur la formulation des demandes d'études, le premier traitant de la formulation des questions qui feront l'objet de l'étude complémentaire (31) et le deuxième du choix de la méthode d'étude la plus adaptée pour répondre à ces questions (32) ;
- un modèle de protocole d'étude (33).

En 2019, EUnetHTA a élaboré un outil spécifique d'évaluation de la qualité des registres accessible en ligne : outil REQueST (*Registry Evaluation and Quality Standards Tool*) (34). Son objectif est de fournir des recommandations pour l'analyse de la qualité des registres en vue de leur utilisation pour l'évaluation des technologies de santé. Dans REQueST, le terme registre est défini comme « un système organisé qui recueille des données et des informations sur un groupe de personnes définies par une maladie, une condition, une exposition ou des services liés à la santé et qui est suivi au fil du temps, servant un objectif scientifique, clinique et/ou de santé publique ». L'outil peut donc s'appliquer à tout recueil continu de données dans une population définie.

Par ailleurs, lors de l'action conjointe en cours, EUnetHTA a mis en place des études pilotes de collaboration pratique entre agences HTA sur la définition conjointe des besoins en matière de données complémentaires à recueillir. Deux types de collaboration existent :

- des collaborations sur un produit de santé spécifique, suite aux informations manquantes identifiées lors de son évaluation. Ces collaborations visent (a) à proposer une question de recherche commune et la définition commune des variables essentielles à collecter dans cette optique, et (b) à préciser les exigences minimales en matière de design et d'aspects statistiques. Les éléments ainsi définis sont utilisés comme base pour la mise en place des études post-inscription au niveau national. Trois collaborations de ce type ont été menées jusqu'en 2020. Les sujets de collaboration (produits) ont été proposés par les agences partenaires du réseau, sur la base des informations manquantes identifiées dans leurs évaluations nationales, et comprenaient : un médicament orphelin en situation d'incertitude à long terme, un médicament dans le cancer du sein métastatique afin de documenter son utilisation en vie réelle, et des dispositifs d'assistance ventriculaire gauche dans l'insuffisance cardiaque terminale ;
- la qualification de registres, qui consiste à évaluer si les données collectées par le registre ainsi que la qualité de leur recueil (évaluée depuis 2019 via REQueST) correspondent aux besoins des agences HTA et peuvent servir comme source adéquate de données en vie réelle.

La collaboration donne lieu à des recommandations élaborées par les agences participantes sur les aspects discutés (variables collectées, qualité des données). Deux projets de ce type ont été menés sur le registre européen de mucoviscidose et sur le registre européen de greffe de moelle osseuse. Ces deux travaux collaboratifs ont été réalisés avec l'implication des détenteurs de registres – et pour le premier, en collaboration avec l'EMA – et font objet des rapports publiés sur le site d'EUnetHTA (35).

Par ailleurs, la HAS reste attentive aux initiatives réglementaires concernant les données en vie réelle dans le domaine du médicament, et notamment au réseau DARWIN EU (36) ainsi qu'à la dernière recommandation de l'Agence européenne du médicament (37), en consultation publique au moment de la rédaction de ce guide, relative aux études en vie réelle.

Table des annexes

Annexe 1.	Références méthodologiques internationales	23
Annexe 2.	Les études en vie réelle descriptives : principes et considérations méthodologiques	25
Annexe 3.	Les études en vie réelle comparatives : principes et considérations méthodologiques	28
Annexe 4.	Les sources de données	34
Annexe 5.	Le <i>Health Data Hub</i> et la base du SNDS	36
Annexe 6.	Mesures centrées sur le patient : l'exemple des <i>Patient Reported Outcome Measure</i>	40

Annexe 1. Références méthodologiques internationales

Thèmes	Sponsor, date	Titre	Résumé du contenu	Liens, auteur
Études observationnelles	ISPE, 2016	Guidelines for good pharmacoepidemiology practice	Recommandations sur la rédaction des protocoles d'études observationnelles, le déroulé de l'étude, la communication des résultats	ISPE, 2015 (38)
Études observationnelles analytiques	STROBE, 2008	The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies	Check-list pour la rédaction des rapports de résultats des études de cohorte, des études cas-témoins et des études transversales	von Elm, et al., 2008 (39)
Études observationnelles en épidémiologie moléculaire	STROBE-ME, 2011	The STrengthening the Reporting of OBservational studies in Epidemiology--Molecular Epidemiology (STROBE-ME): an extension of the STROBE Statement	Recommandations pour les études épidémiologiques concernant les biomarqueurs	Gallo et al., 2011 (40)
Études observationnelles Essais pragmatiques	ENCePP, 2018	Guide on Methodological Standards in Pharmacoepidemiology (Revision 8)	État des lieux régulièrement actualisé des standards méthodologiques en pharmaco-épidémiologie pour la conception, la réalisation et la communication des résultats	ENCePP, 2010 (41)
Études observationnelles	ENCePP, 2018	ENCePP Checklist for Study Protocols (Revision 4)	Check-list pour la rédaction des protocoles d'études observationnelles	ENCePP, 2018 (42)
Études observationnelles	ENCePP, 2018	Code of Conduct (Revision 4)	Recommandations de bonnes pratiques sur le déroulé des études et la communication des résultats dans un objectif de transparence et d'indépendance scientifique	ENCePP, 2018 (29)
Études observationnelles comparatives d'efficacité	ISPOR-ISPE, 2017	Good practices for real-world data studies of treatment and/or comparative effectiveness: Recommendations from the joint ISPOR-ISPE Special Task Force on real-world evidence in health care decision making	Recommandations de bonnes pratiques pour la réalisation et la communication des études observationnelles comparatives d'efficacité	Berger et al., 2017 (43)
Études observationnelles comparatives d'efficacité	AHRQ, 2013	Developing a protocol for observational comparative effectiveness research: A User's Guide	Recommandations pour la réalisation et la publication des études observationnelles comparatives d'efficacité	AHRQ, 2013 (44)
Essais pragmatiques	CONSORT, 2010	CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials	Check-list pour la rédaction des rapports de résultats d'essais cliniques (incluant les essais pragmatiques)	Schulz et al., 2010

				(45)
Études sur bases de données médico-administratives	RECORD, 2015	The REporting of studies Conducted using Observational Routinely-collected health Data	Check-list pour la réalisation d'études sur base de données collectées en routine	Benchimol et al., 2015 (46)
Études sur bases de données médico-administratives	ISPOR-ISPE, 2017	Reporting to Improve Reproducibility and Facilitate Validity Assessment for Healthcare Database Studies V1.0	Recommandations de bonnes pratiques pour la réalisation et la communication des résultats issus d'études sur bases de données médico-administratives	Wang et al. 2017 (47)
Études sur bases de données médico-administratives	Ateliers de Giens, 2019	Requests for post-registration studies (PRS), patients follow-up in actual practice: Changes in the role of databases	Recommandations méthodologiques et procédurales pour la réalisation d'études sur base de données médico-administratives, hospitalières ou de recherche	Berdaï et al., 2018 (48)
Études observationnelles pour l'évaluation des dispositifs médicaux	FDA, 2017	Use of Real-World Evidence to Support Regulatory Decision-Making for Medical Devices.	Recommandations pour l'intégration de données en vie réelle dans les décisions réglementaires américaines (FDA) concernant les dispositifs médicaux	FDA, 2017 (49)
Études pour l'évaluation des dispositifs médicaux (non spécifiquement les études observationnelles)	EUnetHTA, 2015	Therapeutic medical devices guideline	Recommandations pour l'évaluation de l'efficacité clinique comparative des dispositifs médicaux	EUnetHTA, 2015 (50)
Système organisé de collecte de données	EUnetHTA, 2020	The Registry Evaluation and Quality Standards Tool (REQueST®)	Outil d'évaluation de la qualité des systèmes de collecte de données en pratique courante de soin (y compris infrastructure, gouvernance, financement...)	EUnetHTA, 2020 (34)
Revue systématique contenant des études randomisées ou non randomisées	AMSTAR 2, 2017	A critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both	Outil d'évaluation des revues systématiques de la littérature contenant des études randomisées et non randomisées	Shea et al., 2017 (51)
Méta-analyses d'études observationnelles	MOOSE, 2000	Meta-analysis of Observational Studies in Epidemiology	Check-list pour la rédaction de méta-analyses d'études observationnelles	Stroup, et al., 2000 (52)
Études observationnelles	STaRT-RWE, 2020	STaRT-RWE: structured template for planning and reporting on the implementation of real world evidence studies	Recommandations pour la publication et présentation des études observationnelles	Wang et al., 2020 (53)

Annexe 2. Les études en vie réelle descriptives : principes et considérations méthodologiques

Description

Il existe deux grandes familles d'études en vie réelle descriptives en fonction de la nécessité ou non d'un suivi des patients dans le temps :

- **les études transversales**, qui ne comportent qu'un seul point de description des données dans le temps. Elles sont utiles en l'absence de nécessité d'un suivi dans le temps et possibles quand les caractéristiques de la population rejointe ou les conditions de prescription/d'utilisation sont relativement stables dans le temps ;
- **les études longitudinales (particulièrement les études de cohorte)**, au cours desquelles les patients sont suivis dans le temps. Elles sont utiles pour connaître par exemple la durée de traitement, l'observance des patients, les motifs d'instauration et d'arrêt de traitement, etc., ou pour décrire une population rejointe ou des effets indésirables susceptibles d'évoluer dans le temps.

Considérations méthodologiques importantes

Dans une étude en vie réelle descriptive, une collecte primaire ou une utilisation secondaire des données sont possibles (54) (voir annexes 4 et 5), selon la question posée et la disponibilité ou non des informations d'intérêt dans une cohorte ou une base de données existante. Par exemple, les études en vie réelle descriptives sur bases de données médico-administratives ([type SNDS](#)) permettent d'évaluer l'observance par la délivrance du traitement en pharmacie de ville ou à l'hôpital (si inscrit sur la liste en sus). En revanche, elles ne permettent pas d'évaluer la consommation réelle du produit de santé, l'expérience du patient ou son état de santé ou de connaître la surveillance clinique faite par le médecin, hors actes de biologie ou de radiologie. S'agissant d'analyses portant sur des DM implantables, si plusieurs implants avec le même code LPP ont été utilisés successivement pour traiter différentes localisations anatomiques, la possibilité d'un chaînage des informations avec l'implant d'intérêt doit être vérifiée pour qu'une étude à partir des bases de données d'hospitalisation puisse être envisagée (par exemple, dans certains cas, le côté opéré est individualisé ; dans d'autres cas et notamment lorsqu'il s'agit de traitements successifs de plusieurs niveaux anatomiques, l'information n'est pas disponible, ce qui peut être un frein à une démarche d'évaluation sur les bases de données).

Dans le cas de la réalisation *ad hoc* d'une étude, certaines recommandations méthodologiques décrites ci-dessous doivent être prises en compte dès la planification de l'étude. En cas d'utilisation secondaire de données, le respect de ces mêmes éléments doit être vérifié.

Sélection de la population d'étude

La méthode de sélection des centres et des patients (échantillonnage) inclus dans l'étude en vie réelle doit faire l'objet d'une attention particulière car l'objectif sera de décrire, de façon la plus proche possible de la réalité, les professionnels et patients utilisant le produit de santé en pratique courante. En amont de la réalisation de l'étude, plusieurs techniques permettent de réduire le risque de biais de sélection dans l'étude :

- l'identification des lieux de soins reflétant au mieux les lieux où est prise en charge la population d'intérêt (centres de soins primaires, secondaires ou tertiaires ?) et éventuellement leur proportion respective, ou à l'inverse la file active de ces différents types de prescripteurs vis-à-vis de

la pathologie d'intérêt. Par exemple, si les patients souffrant d'une pathologie particulière sont pris en charge à 20 % en centre hospitalier et à 80 % par des médecins généralistes, il faudra inclure suffisamment de médecins généralistes pour retrouver ces pourcentages dans l'étude ;

- l'identification des prescripteurs ou implanteurs possibles du produit de santé en vie réelle (caractérisation des praticiens, discipline et cadre d'exercice). Par exemple, si les patients souffrant d'une pathologie particulière sont suivis à 50 % par des endocrinologues et 50 % par des rhumatologues, les médecins investigateurs devront refléter cette diversité.

Par ailleurs, pour pallier un biais de sélection lié à l'investigateur et afin que l'investigateur inclue tous les patients éligibles indépendamment de sa préférence, une inclusion consécutive des patients est nécessaire.

En début d'analyse statistique, il est recommandé d'évaluer la présence d'un biais de sélection en comparant les caractéristiques des centres et des professionnels ayant refusé de participer à l'étude, à celles des professionnels inclus. Concernant les patients, il est parfois possible de collecter quelques données démographiques (âge, sexe) des patients ayant refusé l'étude, pour les comparer à celles de patients inclus. Le recueil des caractéristiques des patients refusant l'inclusion doit être autant que possible documenté.

Dans certains cas particuliers comme dans le cadre d'une maladie rare, l'exhaustivité est importante et l'étude doit idéalement inclure tous les patients atteints de la pathologie. Pour ce faire, il faudra vérifier que tous les services susceptibles de prendre en charge les patients atteints de la maladie ont été sollicités et que tous les patients de ces services ont été inclus dans l'étude. Pour ce faire, plusieurs sources d'information peuvent être nécessaires.

Il est à noter qu'un biais de sélection ne peut jamais être totalement exclu dans une étude nécessitant un consentement.

Données collectées

→ Caractère observationnel de l'étude

Dans le cadre des études descriptives *ad hoc*, le promoteur doit s'assurer que la mise en place de l'étude ne va pas modifier les habitudes et choix du prescripteur. La collecte des données (par exemple via un cahier d'observation – *Case Report Form*) doit permettre de collecter des données pertinentes pour l'étude sans pour autant modifier les pratiques médicales.

Par ailleurs, il pourrait paraître suspect qu'un critère d'inclusion dans l'étude soit l'initiation de, ou un switch vers, l'intervention d'intérêt pour une proportion importante de patients. Ceci reviendrait, par le design même de l'étude, à encourager les prescripteurs à initier cette intervention pour pouvoir inclure des patients dans l'étude (*seeding trial*) (55).

→ Niveau de détail et granularité des données

L'intérêt d'une étude en vie réelle étant souvent de fournir des informations non disponibles dans les essais cliniques, une attention particulière doit être portée sur l'étendue et l'exhaustivité des informations collectées (caractéristiques des centres de soins, des prescripteurs, des patients, des modalités de prescription et d'utilisation du produit de santé).

Pour s'assurer que les données pertinentes vont être collectées, il est utile de demander un avis auprès d'experts cliniques les plus à même de connaître les variables d'intérêt permettant de répondre aux objectifs de l'étude. En cas d'utilisation secondaire de données, il faut s'assurer que la cohorte (registre) ou la base de données médico-administratives contient les données pertinentes pour l'étude.

➔ Qualité des données

Bien que le suivi des patients soit observationnel dans ces études, il est impératif de limiter le nombre de patients perdus de vue et donc le risque de données manquantes, notamment vis-à-vis de la mesure de l'évènement d'intérêt principal (statut final du patient). Lorsqu'un patient est perdu de vue et que son statut final est inconnu, il est impossible de le catégoriser comme ayant ou n'ayant pas eu l'évènement d'intérêt principal. Un nombre trop important de patients perdus de vue rend les résultats d'une étude peu valides, voire non interprétables.

Le risque de données manquantes sur le statut final du patient est d'autant plus grand que la durée de suivi est longue.

Pour minimiser le risque de perdus de vue au moment de la réalisation de l'étude, il peut être envisagé de motiver les patients à participer à l'étude sur toute sa durée, par exemple en prévoyant des relances en cours d'étude, au moyen d'outils numériques, ou de toute autre manière qui doit être décrite dans le protocole d'étude.

Si l'évènement d'intérêt principal est un décès, le statut vital des patients perdus de vue doit toujours être scrupuleusement recherché auprès de l'état civil de la commune de naissance du patient ou par recoupement avec le registre national des causes de décès français (Inserm – Centre d'épidémiologie sur les causes médicales de décès – CépiDc).

Enfin, lors de l'analyse statistique, le nombre et pourcentage de patients perdus de vue devront être décrits et les caractéristiques de ces patients devront être comparés à celles des patients non perdus de vue.

Annexe 3. Les études en vie réelle comparatives : principes et considérations méthodologiques

Description

Études de cohorte

Le suivi de la cohorte permet de réaliser des analyses comparatives afin de déterminer si les patients « exposés » au produit de santé, comparativement aux patients « non exposés » (ou exposés au(x) comparateur(s)), présentent un risque relatif inférieur, égal ou supérieur d'avoir l'évènement considéré, avec possibilité d'une inférence causale du lien (si contrôle des biais appropriés, notamment les facteurs de confusion) entre ces deux éléments.

Études cas-témoins

Elles reposent sur la constitution d'un groupe de patients « cas » présentant une pathologie étudiée (ou un évènement en lien avec une pathologie : rechute, hospitalisation, etc.) et d'un groupe de patients « témoins » n'ayant pas cette pathologie. Le principe des études cas-témoins est de comparer la fréquence de l'exposition des patients dans les deux groupes. Ce type d'étude est adapté dans les cas où la pathologie étudiée est rare et l'exposition fréquente. Cependant, la constitution du groupe témoin est un processus difficile : les témoins doivent être identiques aux cas, sauf pour la pathologie en question, et doivent provenir de la même cohorte théorique que les cas. Ce type d'étude permet de calculer un *odds ratio* (OR) qui peut approximer le risque relatif de la survenue de l'évènement d'intérêt chez les sujets exposés par rapport aux sujets non exposés, sous certaines conditions. En effet, cette approximation est possible quand la pathologie est peu fréquente (< 10 %) et la valeur du risque relatif n'est pas trop élevée (< 5 %). Il existe quelques cas particuliers d'études cas-témoins.

– Les études cas-témoins nichées dans une cohorte

Les cas et les témoins sont identifiés dans une même cohorte définie antérieurement à l'étude. Ces études permettent de réduire les biais de sélection lors du choix des témoins, ainsi que les biais de mémorisation, inhérents aux études cas-témoins. L'inconvénient majeur de ce type d'étude est la non-représentativité potentielle de la population des sujets indemnes (à partir de laquelle seront tirés au sort les témoins) par rapport à la cohorte initiale, du fait de l'attrition des patients depuis le début de la cohorte (survenue de décès ou perdus de vue).

– Les études case cross-over

Ce type d'étude est pertinent lorsque l'hypothèse peut être faite d'un délai très court entre l'exposition et l'évènement, comme l'effet à très court terme d'un produit de santé. Seuls des patients ayant eu l'évènement d'intérêt sont inclus. La période précédant directement l'évènement est dite période « à risque » (dans laquelle on mesure l'exposition ou non de chaque patient) et la période précédant l'évènement de plus loin est dite période de « référence ». Chaque patient est considéré comme son propre « témoin ». Le rapport d'exposition dans la période à risque est comparé au rapport d'exposition dans la période de référence.

Considérations méthodologiques importantes

Les études en vie réelle comparatives sont une source potentielle de connaissance sur l'effet des traitements dans des conditions moins contrôlées que celles des essais cliniques classiques (43, 56), notamment par rapport à des comparateurs qui n'ont pas été évalués dans les essais cliniques classiques.

Elles offrent la possibilité de répondre à des questions que les essais cliniques classiques n'explorent pas, notamment sur les éléments contextuels qui peuvent interagir avec l'effet propre du traitement et en modifier l'efficacité en vie réelle. Néanmoins, la réalisation du protocole, les analyses et l'interprétation des résultats doivent se faire au regard d'erreurs systématiques possibles (biais) liées à la sélection des patients, à la mesure de l'exposition ou de l'évènement, ou à la présence de facteurs de confusion dans certaines associations que l'on souhaite mettre en évidence (57, 58).

→ **Biais de sélection liés aux patients inclus**

Outre le risque de biais de sélection décrit dans le paragraphe relatif aux études en vie réelle descriptives, d'autres biais de sélection liés aux patients inclus existent, notamment :

- le **biais de déplétion des susceptibles**, qui est un biais lié aux traitements passés des patients éligibles à l'étude et qui concourt à surestimer le bénéfice d'un traitement ou à en sous-estimer le risque. En effet, les patients ayant pris le traitement d'intérêt par le passé et l'ayant arrêté (par exemple pour cause d'effet indésirable ou d'un manque d'efficacité) ne seront pas comptabilisés comme « exposés » au traitement en question au début de l'étude. En d'autres termes, la cohorte théorique des patients exposés au traitement d'intérêt s'est « vidée », au fil du temps, des patients pour qui le traitement n'était pas adapté. Ce risque de biais rend pertinent le recours aux patients incidents (*new-user design*), qui consiste à n'inclure que des patients prenant le médicament ou utilisant le DM d'intérêt pour la première fois de leur vie (59) ;
- le **healthy user effect**, qui correspond à une corrélation positive fréquemment observée entre la qualité ou quantité des comportements d'hygiène de vie et la propension à accepter ou être observant vis-à-vis d'un traitement. Ces comportements améliorent leur pronostic, sans lien avec le traitement lui-même. La prise en compte de ces facteurs comportementaux est donc importante dans une étude observationnelle.

→ **Biais de mesure et de classification**

Dans une étude de cohorte, le **biais de temps immortel** correspond à la prise en compte dans les analyses d'un temps « d'exposition à un traitement » qui inclurait en réalité la période où les patients du groupe « exposé » n'avaient pas encore ce traitement, c'est-à-dire la période entre l'évènement ayant conduit au diagnostic (exemple névrite optique rétrobulbaire dans la sclérose en plaques) et l'initiation effective du traitement (anti-inflammatoires). Une des mesures pour pallier ce biais est de réaliser les analyses selon les « personnes-temps ». Cette période ne doit pas être prise en compte dans le nombre de « personnes-temps » exposées mais ajoutée au nombre de « personnes-temps » non exposées (60, 61).

Dans les études cas-témoins, un **biais de mémorisation** correspond au fait que les personnes ayant eu l'évènement d'intérêt ou leurs proches sont susceptibles de mieux se souvenir de l'exposition collectée *a posteriori*, que celles n'ayant pas eu l'évènement.

→ **Biais d'attrition et données manquantes**

Dans une étude observationnelle, les stratégies mises en place pour favoriser la participation des patients jusqu'au bout de l'étude ne sont pas toujours optimales. Le risque de perdus de vue et de données manquantes pour la mesure d'intérêt principale est donc plus important que dans un essai clinique classique. De même, les investigateurs peuvent être moins précis ou réguliers dans la collecte des données – risque qu'il faudra minimiser.

La présence d'un biais de mesure lié à ces données manquantes dépend de leur fréquence et de la raison de la perte d'information. Lorsque la donnée sur la mesure d'intérêt principale arrive de façon totalement aléatoire et sans lien avec les facteurs liés au patient ou à l'effet du produit de santé, la donnée manquante est dite *Missing Completely At Random*, situation extrêmement rare. Plus généralement, la raison de la donnée manquante peut être expliquée par les caractéristiques du patient (*Missing At Random*, MAR), voire également expliquée par la survenue de l'évènement d'intérêt. Dans ce dernier cas, les données manquantes sont informatives et dites *Not Missing At Random* (NMAR).

Une complication s'ajoute si la fréquence des données manquantes n'est pas la même dans les deux groupes d'intervention.

Les moyens de diminuer le nombre de données manquantes incluent :

- pendant le déroulé de l'étude : la réduction de la pénibilité de participation à l'étude pour les patients afin de diminuer l'attrition ;
- lors des analyses (62) : si les données manquantes ne sont pas trop fréquentes et si les données sont supposées être MAR, l'imputation multiple des données manquantes est possible ; pour des données manquantes de type NMAR, les techniques d'imputation sont très limitées ; des analyses de sensibilité peuvent être utiles. Une analyse sur cas complets est en revanche non recommandée, sauf pour les analyses de sensibilité, en considérant alternativement les patients perdus de vue comme des succès ou des échecs de traitement.

→ **Biais d'indication et autres effets de confusion**

Dans une étude en vie réelle, et contrairement aux essais cliniques classiques, la réalisation de l'étude ne doit pas modifier – ou le moins possible – le déroulé habituel des soins et le choix du produit de santé prescrit (63). En pratique courante de soins, le choix réalisé par les prescripteurs entre différents dispositifs de soins possibles se fait en fonction de facteurs liés au patient (critères pronostiques (64), comorbidités (65)), aux traitements (effets indésirables, traitements passés (66)), voire aux habitudes du prescripteur. Par conséquent, et du fait de la non-randomisation, ces facteurs ne sont pas identiques chez les patients recevant le produit de santé d'intérêt et ceux recevant les alternatives (comparateurs). Par ailleurs, ces différences tendent à varier avec le temps, ajoutant un niveau supplémentaire de complexité aux analyses.

Dès lors, si l'objectif de l'étude est de réaliser une inférence causale sur le lien entre « l'exposition » à un produit de santé (par rapport à un autre) et la survenue d'un évènement d'intérêt, il faut prendre en considération ces effets de confusion (biais d'indication et *channeling bias*) pour mesurer correctement la taille d'effet et pouvoir attribuer cette différence d'effet au seul traitement. Lorsque le poids des facteurs de confusion varie dans le temps, on parle de biais d'indication évoluant avec le temps (*time-varying channeling bias*) (67, 68).

L'identification des facteurs de confusion potentiels doit se faire en amont lors de la planification de l'étude grâce à une expertise clinique et/ou une revue de la littérature. Il faut garder à l'esprit que les facteurs de confusion ne sont pas toujours tous observables ou connus et que leur prise en compte est donc fréquemment imparfaite. Néanmoins, il est important que les facteurs de confusion observables soient bien collectés lors de l'étude, ou présents dans la base de données utilisée.

Au niveau des analyses statistiques, les méthodes de prise en compte des facteurs de confusion incluent, mais ne se limitent pas, à l'utilisation :

- d'un modèle multivarié avec appariement et/ou ajustement sur les facteurs de confusion ;

- d'un modèle multivarié avec appariement et/ou ajustement sur les scores de propension simples ou à haute dimension (69) ;
- d'un modèle marginal structural avec pondération sur la probabilité inverse de traitement (*inverse-probability of treatment weighing*) (70) ;
- de variables instrumentales (*instrumental variables*) ;
- de méthode en doubles différences ;
- de régressions sur discontinuités.

Il est à noter que même si l'étude est comparative, les analyses descriptives sont une étape préalable indispensable à toute analyse inférentielle. La description des différents groupes apporte des éléments importants. De même, il peut être instructif de visualiser la distribution des scores de propension et le degré de superposition des courbes de distribution de ces scores de propension (stratégie d'intérêt vs comparateur) qui indique le degré global de similitude des patients des deux groupes vis-à-vis des facteurs de confusion observés, ainsi que la possibilité ou pas d'utiliser la méthode des scores de propension.

Essais pragmatiques

Description

Les essais pragmatiques sont des essais cliniques randomisés où les éléments de design traditionnels des essais cliniques sont rendus plus flexibles, afin de refléter davantage les pratiques courantes (71). La notion d'approche pragmatique d'un essai clinique, par opposition à une approche expérimentale, a été introduite en 1967 par Schwartz et Lellouch (72). Le propos des auteurs était qu'un essai clinique peut répondre à différentes questions qui nécessitent des approches méthodologiques différentes. L'approche expérimentale sert à démontrer l'efficacité pharmacologique d'un médicament, en ayant contrôlé tous les autres éléments pouvant expliquer les résultats observés – d'où le cadre expérimental strict. L'approche pragmatique vise, quant à elle, à évaluer l'efficacité du traitement dans les conditions usuelles de prise en charge (*effectiveness*). Bien que les essais pragmatiques soient plus proches de la pratique courante, ils reflètent une expérimentation et ne sont donc pas considérés comme des études observationnelles. Les éléments de design des essais pragmatiques ont les caractéristiques suivantes (73, 74) :

- les centres recrutant les patients doivent refléter les lieux habituels de soins ;
- les critères d'éligibilité des patients doivent être larges et non strictement ceux de l'AMM ;
- l'unité de randomisation peut être le patient ou le lieu de soins (*cluster-randomized trials*) ;
- le comparateur doit refléter les alternatives de soins et les traitements habituellement prescrits dans cette indication ;
- la collecte des données doit être aussi proche que possible de la prise en charge classique de la pathologie concernée et minimiser la pénibilité pour le prescripteur et tant que faire se peut, utiliser des dossiers médicaux informatisés (75) ;
- la mesure d'intérêt principale doit être la plus pertinente possible pour le prescripteur, par exemple celle principalement utilisée pour adapter la thérapeutique, ou pour le patient, par exemple un critère de type PRO.

Les autres éléments relatifs au design de l'essai clinique pragmatique doivent également être choisis pour refléter au mieux les pratiques courantes (monitoring de l'intervention, fréquence du suivi, surveillance du traitement, etc.). Le design le plus simple d'un essai pragmatique consiste à randomiser les participants en deux groupes d'intervention qui seront comparés. Il existe d'autres types de designs.

- **Cohort-multiple randomized trial** : les patients d'une cohorte existante sont randomisés, soit pour changer de traitement (groupe « intervention d'intérêt »), soit pour rester avec leur traitement (groupe comparateur). Seuls les patients randomisés dans le premier groupe sont sollicités pour consentement à l'inclusion dans l'essai, les autres étant déjà suivis de façon observationnelle. Ce design permet de diminuer la pénibilité du recrutement et du consentement pour au moins la moitié des patients et diminuer le coût de l'étude. Une limite importante liée à ce design est la possibilité de refus dans le groupe « intervention d'intérêt » et pas dans l'autre, ce qui a pour conséquence un potentiel biais de sélection et une perte de puissances statistique et méthodologique.
- **Cluster-randomized trial** : l'unité de randomisation est le centre de soins et non le patient. Les patients inclus dans l'essai recevront le traitement pour lequel le centre qui s'occupe d'eux a été randomisé. Ce design permet de réduire le biais de contamination, dans la mesure où l'aveugle n'est pas souhaité pour être plus proche de la réalité, en conformité avec les principes des essais pragmatiques (71). Les limites importantes de ce design incluent : (a) un risque de biais de sélection par le prescripteur lié au fait qu'il connaît le produit de santé qui est prévu d'être initié chez ses patients (76) ; (b) une perte de puissance statistique et (c) la possibilité d'un effet cluster qu'il faut prendre en considération via les analyses d'interaction.

Considérations importantes

La modification de paramètres de design visant à rendre un essai clinique plus pragmatique peut induire certains biais qu'il faut anticiper et minimiser dès la planification de l'étude et prendre en compte au moment des analyses statistiques. Il s'agit des biais de sélection liés à la préférence des patients et des prescripteurs, hétérogénéité des patients et de l'effet du traitement avec risque de dilution de l'effet mesuré, biais de mesure si l'évènement d'intérêt principal est subjectif, etc. (64-66, 77, 78). De façon plus importante, la question posée et à laquelle doit répondre l'essai pragmatique doit être précisément formulée, afin de juger de ces risques au regard de ce que l'on souhaite précisément mesurer (*estimand*).

Modélisation prédictive

La modélisation prédictive vise à extrapoler des observations du présent à un autre contexte (le futur, une autre population, un autre pays). Les essais cliniques sont réalisés dans un contexte particulier et restrictif vis-à-vis des pratiques courantes, par exemple :

- ils sont limités dans le temps ;
- ils ne sont pas réalisés dans l'ensemble des pays où le produit de santé sera prescrit ;
- la population analysée n'est pas identique à la population cible ;
- la population analysée n'est pas forcément celle dans laquelle le produit de santé aura le meilleur impact.

Des techniques d'intelligence artificielle (exemples, *machine learning* comme les réseaux de neurones), ou d'autres méthodes de modélisation statistique, peuvent être utilisées pour compléter les résultats d'un essai clinique ou d'autres types d'étude, afin de prédire ce que serait l'effet du produit de santé dans le « nouveau » contexte d'intérêt (effet à long terme, effet en France, effet dans la population cible, etc.).

Dans l'industrie pharmaceutique, ces méthodes sont parfois utilisées pour prédire les résultats de l'essai de phase 3 (79, 80) à partir des données disponibles préalablement. La FDA puis l'EMA

encouragent les réflexions sur le sujet des simulations et des essais cliniques *in silico* pour optimiser le développement des médicaments et accompagner leur mise sur le marché.

À noter :

- il faut distinguer « données simulées » (les données n'existent pas, elles sont générées par informatique) et les études de simulation (peu importent les données, différents scénarios sont simulés, pour optimiser une méthode de prédiction et analyser sa sensibilité à différentes contraintes) ;
- ce que l'on appelle *big data*, qui correspond à des bases de données de taille très importante, de l'ordre de 100 millions de lignes et colonnes, voire plus.

Principes généraux

Ces méthodes (algorithmes, *machine learning*, réseaux de neurones, etc.) peuvent utiliser des statistiques fréquentistes ou bayésiennes. Les modèles peuvent être paramétrés de façon « supervisée » (c'est-à-dire que la variable à prédire est définie *a priori*, le modèle cherchant à expliquer cette variable) ou « non supervisée » (la méthode classe des sujets sans choix *a priori* de la variable à prédire).

Pour certaines techniques de *machine learning*, ce qui importe, c'est la capacité du modèle à prédire correctement ce qui se passe dans le « nouveau » contexte d'intérêt, en utilisant des indicateurs de puissance prédictive. En revanche, l'interprétation des coefficients d'un modèle (par exemple les coefficients de régression) et le sens clinique que l'on donne à une association observée peuvent être de moindre importance.

Considérations méthodologiques

Le *machine learning* ou toute autre technique de modélisation statistique prédictive se réalise en plusieurs phases successives :

- la phase d'apprentissage (*training*), qui se réalise sur tout ou partie de la base de données dite d'apprentissage (*training set* : base utilisée pour spécifier le modèle prédictif) : plusieurs modèles peuvent être spécifiés ;
- la validation interne du modèle en utilisant le *training set* ;
- la phase de test ou validation externe, qui se réalise sur tout ou autre partie de la base de données (*test set*) ou sur une ou plusieurs bases de données externes, pour apprécier la capacité du modèle à prédire des situations autres que celles utilisées pour la phase d'apprentissage, permettant ainsi d'apprécier d'éventuels problèmes d'*overfitting* (cf. ci-dessous), lors de cette phase, la puissance prédictive du modèle est évaluée ;
- la phase de réapprentissage.

Risques et limites méthodologiques

- Le surajustement (*overfitting*) : la puissance prédictive du modèle sur le *training set* semble élevée, voire quasi parfaite (*overfitting*), c'est-à-dire que le modèle présente une bonne, voire excellente validité interne, mais sa validité externe (capacité à prédire les événements du *test set*) est faible. Ce modèle n'est donc pas efficace. De nombreuses approches statistiques peuvent être utilisées pour spécifier des modèles tout en prenant en compte le risque d'*overfitting*.
- L'extrapolation (*coverage* de l'entraînement) : le modèle prédictif testé sur une autre source de données (autre pays, autre période) peut s'avérer insatisfaisant ; le modèle doit donc constamment être adapté.

Annexe 4. Les sources de données

Données primaires

La collecte primaire de données signifie que la collecte des données de l'étude a été planifiée et réalisée dans le cadre d'une étude *ad hoc*. Ces données peuvent être collectées par les médecins investigateurs, des attachés de recherche clinique à partir de dossiers médicaux, ou des patients (auto-questionnaires).

Les données collectées forment une cohorte, ou un registre lorsque le recueil des données est continu et exhaustif dans une population géographiquement définie (81).

Une collecte *ad hoc* de données peut s'avérer nécessaire si aucune source de données déjà existantes ne permet de répondre aux objectifs spécifiques de l'étude à mener.

Données secondaires

L'utilisation secondaire de données signifie que les données utilisées pour l'étude avaient déjà été collectées pour un autre but (82).

Les principales sources de données secondaires sont :

- les données médico-administratives du SNDS ;
- les données issues de la production de soins en routine ;
- les systèmes d'information hospitaliers, en particulier les entrepôts hospitaliers de données ;
- les systèmes d'information publics ou privés en médecine de ville, notamment basés sur les logiciels d'aide à la prescription ;
- les données des autres acteurs de la production de soins en ville (pharmacies d'officine, laboratoires de biologie et centres d'imagerie) ;
- les données issues de registres et cohortes ;
- les données produites par les patients, soit par le biais de dispositifs médicaux connectés d'automesure (rythme cardiaque, glycémie...), soit par la saisie directe par le patient d'informations dans des applications.

Depuis maintenant plusieurs années, l'utilisation de données secondaires a pris une place importante dans la réalisation des études en vie réelle, à la suite de l'ouverture de l'accès au SNDS et l'organisation progressive d'une utilisation secondaire des données de cohortes académiques. L'organisation de la collecte et l'exploitation de nouvelles données de santé sont des champs en pleine expansion et leur utilisation pour des recueils de données en vie réelle devrait continuer à s'étendre (48).

Considérations relatives au choix d'une source de données secondaires

Lors du choix d'une source de données pour la réalisation d'une étude en vie réelle, il est important de vérifier les points suivants (44, 48).

- La source de données envisagée permet-elle de répondre à l'objectif de l'étude en vie réelle ? Notamment, toutes les mesures d'intérêt sont-elles collectées dans la base de données ? Autrement dit, il faut s'assurer que les variables disponibles dans la base de données permettent d'identifier précisément : la maladie ou un sous-type, les éléments cliniques pertinents, des mesures de qualité de vie, les variables mesurant les facteurs de confusion potentiels, etc. Ce type de données n'est pas présent dans les bases du SNDS.
- Les données recherchées sont-elles accessibles indirectement par l'intermédiaire d'algorithmes ? Ces derniers sont-ils validés ?
- Les données recherchées sont-elles accessibles et de bonne qualité ? Les données collectées doivent faire l'objet d'une évaluation de leur qualité, incluant des critères métrologiques tels que

leur exactitude, mais impliquant aussi l'évaluation des processus de saisie, la normalisation ou non de leurs valeurs (y compris les référentiels et dictionnaires et leurs versions successives), les processus de contrôle (sur fichiers bruts et préparés en vue d'exploitation pour l'étude en vie réelle), de correction, de reclassification, de traitement des données manquantes, etc.

- La représentativité des données est-elle assurée ?
- La temporalité de l'exposition au produit de santé et de la collecte des données d'intérêt est-elle compatible avec la chronologie de l'étude en vie réelle ?
- La transparence vis-à-vis de la source de financement de la source de données est-elle suffisante ; les conflits d'intérêts possibles sont-ils définis ?

Utilisation secondaire de cohortes et registres académiques

Les cohortes ou registres académiques déjà constitués représentent une source de données potentiellement très intéressante pour la réalisation d'études en vie réelle. Plusieurs cohortes d'envergure nationale ont d'ores et déjà été utilisées dans des évaluations de médicaments par la CT : la cohorte HEPATHER⁶ dans les avis de réévaluation des antiviraux d'action directe dans le traitement de l'hépatite C, la cohorte OFSEP⁷ pour la réévaluation des médicaments de la sclérose en plaques, la cohorte CKD Rein⁸ pour l'évaluation du tolvaptan dans la polykystose rénale, la cohorte RaDiCo⁹ dans la fibrose pulmonaire idiopathique. Le portail Épidémiologie-France propose un catalogue en ligne des principales bases de données en santé et études de source française. Le catalogue de la plateforme française *Health Data Hub* devrait s'enrichir prochainement et proposer plusieurs bases de données de santé (voir ci-dessous).

Les atouts potentiels des études de cohortes et registres académiques dans le cadre des études en vie réelle sont :

- l'adhésion des professionnels de santé qui permet de limiter le risque de faible participation ;
- la richesse des données cliniques et biologiques permettant de répondre à des questions spécifiques auxquelles il ne serait pas possible de répondre avec les données de remboursement ou d'hospitalisation par exemple ;
- la collecte d'informations sur plusieurs années permettant l'évaluation de l'efficacité et des effets indésirables à long terme des produits de santé ;
- la possibilité d'étudier l'ensemble des traitements d'une indication donnée au contraire des études centrées sur le suivi d'un médicament ou d'un dispositif médical ;
- la qualité scientifique des cohortes d'envergure nationale.

Les limites sont notamment :

- l'absence de données importantes pour le suivi des médicaments et dispositifs médicaux du fait qu'elles n'ont pas été conçues pour répondre à l'objectif de l'étude en vie réelle ;
- le manque de représentativité pour les cohortes constituées à partir de centres sélectionnés ;
- l'accès aux données nécessite la mise en place d'un partenariat entre l'équipe académique et les laboratoires ; ceci est un processus long, qui doit être anticipé et organisé ;
- la qualité de la collecte des données (ex. : risque de données manquantes et de biais de sélection) et qui varie d'une cohorte à l'autre.

⁶ ANRS CO22 HEPATHER, Options thérapeutiques au cours des hépatites B et C : une cohorte nationale française.

⁷ Observatoire français de la sclérose en plaques.

⁸ *Chronic Kidney Disease* – Réseau épidémiologie et information en néphrologie.

⁹ *Rare Disease Cohorts*.

Annexe 5. Le *Health Data Hub* et la base du SNDS

Le *Health Data Hub* est la plateforme française des données de santé (<https://www.health-data-hub.fr/>) créée par la loi du 24 juillet 2019 relative à l'organisation et la transformation du système de santé. Il s'agit d'une structure publique permettant à des porteurs de projets d'accéder à des données de santé dans un cadre très sécurisé, pour qu'ils contribuent à trouver des solutions qui améliorent la santé des personnes. Des procédures très précises encadrent ces accès, et surtout ces données, qui ne contiennent ni les noms, ni les prénoms, ni les numéros de sécurité sociale des personnes. L'objectif du *Health Data Hub* est donc que les données de santé, qui représentent un véritable patrimoine collectif, puissent être utilisées au bénéfice de la société. Ces données sont notamment celles de la base principale du système national des données de santé (SNDS), mais également des établissements de santé, de Santé publique France, etc. Ces autres sources sont réunies dans un « catalogue », c'est-à-dire une collection de bases de données non exhaustives qui sera construite de manière progressive et itérative, en partenariat avec les responsables de la collecte des données concernées. La réplication de ces bases d'intérêt au niveau d'un même dispositif permettra de mutualiser les investissements humains, technologiques et financiers en vue de :

- enrichir de manière pérenne chacune des bases avec les données médico-administratives (i.e. la base principale du SNDS) ;
- agréger, chaîner, extraire et mettre à disposition des données ciblées via la plateforme technologique du *Health Data Hub* pour les utilisateurs habilités ;
- donner une plus grande visibilité de la base avec un potentiel impact sur le taux de réutilisation des données ;
- offrir un hébergement sécurisé et conforme au RGPD.

Ce catalogue, c'est la possibilité de mener des recherches d'intérêt public avec de forts impacts pour la société. La liste des flux de la base principale et des bases composant le catalogue est portée par un arrêté (à paraître au moment de la rédaction de ce guide).

La consultation des équipes du *Health Data Hub* est recommandée pour obtenir un appui technique et réglementaire concernant l'accès au catalogue de données et/ou la conception d'une base de données en vue de son hébergement et d'un possible chaînage au SNDS.

Focus sur le système national des données de santé

Le SNDS a été créé par la loi de modernisation de notre système de santé de janvier 2016. Il s'agit d'un entrepôt de bases de données médico-administratives qui permet de chaîner plusieurs sources de données de santé issues :

- du système national d'information inter-régimes de l'Assurance maladie (Sniiram), base de l'Assurance maladie créée en 1999 par la loi de financement de la sécurité sociale ;
- de la base du programme de médicalisation des systèmes d'information (PMSI) contenant les données des hôpitaux ;
- de la base du CépiDc de l'Inserm recensant les causes médicales de décès ;
- et des données médico-sociales du système d'information des maisons départementales des personnes handicapées (MDPH).

Le SNDS se décline ainsi en différentes bases de données individuelles et comprend également un échantillon représentatif au 1/97 (échantillon généraliste des bénéficiaires (EGB)) de la population protégée, ainsi que des bases thématiques avec des données agrégées (Datamarts BO).

Les finalités et le périmètre du SNDS ont évolué au fil des années. Initialement mis à disposition des agences de l'État, l'accès à ces bases de données a été progressivement élargi à la communauté

scientifique, sous réserve d'une autorisation d'accès délivrée par projet. Les accès temporaires à des fins de recherche sont règlementés. Dans le cas général, il est nécessaire d'obtenir une autorisation auprès de la Cnil pour traiter des données personnelles de santé. En amont de sa décision, deux institutions sont mobilisées dans le cadre de la procédure :

- le *Health Data Hub*, guichet unique des projets mobilisant des données de santé ;
- le Comité éthique et scientifique pour les recherches, les études et les évaluations dans le domaine de la santé (CESREES), comité indépendant en charge de rendre un avis notamment sur la méthodologie retenue, sur la nécessité du recours à des données à caractère personnel, sur la pertinence de celles-ci par rapport à la finalité du traitement et, s'il y a lieu, sur la qualité scientifique du projet.

Il existe également des procédures d'accès dites simplifiées, dans le cadre desquelles l'autorisation de la Cnil n'est pas requise, c'est notamment le cas si le traitement envisagé est conforme à une méthodologie de référence (dite MR).

La base principale du SNDS regroupe les données des patients affiliés à différents régimes de l'Assurance maladie et peut être considérée comme quasi exhaustive des données de la population française. Après un processus de pseudonymisation (83), il est ainsi possible d'accéder à des données individuelles des bénéficiaires (année de naissance, sexe, département de résidence, date et commune de décès, la notion de couverture maladie universelle complémentaire, médecin traitant), leur inscription sur la liste des affections de longue durée (ALD), des données concernant les séjours hospitaliers issues des données du programme de médicalisation des systèmes d'information, PMSI (médecine, chirurgie, obstétrique, psychiatrie, soins de suite et de réadaptation, hospitalisation à domicile avec des diagnostics médicaux dits principaux, reliés et associés, les dates de séjour, les procédures thérapeutiques ou d'imagerie réalisées) et tous les soins remboursés en ville (consultations médicales, actes infirmiers, actes de biologie ou d'imagerie, délivrance de médicaments ou de dispositifs médicaux remboursés : données de consommation inter-régimes, DCIR).

Intérêts et limites

Le principal atout de la base principale du SNDS est la quasi-exhaustivité des données de la population française. La granularité de l'information est importante. Par exemple, parmi les données disponibles sur les médicaments remboursés par l'Assurance maladie figurent : le nom de la molécule, le nom de la spécialité, le dosage et le nombre d'unités thérapeutiques, galéniques. En revanche, il convient de préciser qu'en identifiant des dates de délivrance de médicaments ou de dispositifs médicaux remboursés, l'exposition réelle au produit de santé ne peut être qu'approximée. L'observance thérapeutique n'est en effet pas directement observable et les seuls indicateurs disponibles sont :

- la régularité des dates de délivrance ;
- les quantités de produits de santé délivrées ;
- la durée entre la première et la dernière date de délivrance observée.

De même, la notion de surveillance clinique ne pourra tirer profit que des dates de réalisation des actes de biologie ou de radiologie, sachant que leurs résultats ne sont pas disponibles.

Toutefois, dans la mesure où le SNDS est inter-régimes, les études réalisées sur ces bases de données peuvent raisonnablement revendiquer leur quasi-exhaustivité et représentativité de la population française. Il faut cependant garder à l'esprit que l'alimentation des bases n'est pas instantanée avec des périodicités différentes selon les sources de données. Par ailleurs, le temps incompressible de l'étape réglementaire puis de la mise à disposition des données doit être pris en compte dans la réalisation d'une étude.

Un deuxième avantage du SNDS est l'absence de biais d'attrition et un taux de données manquantes faible. La problématique des perdus de vue est cruciale dans les études observationnelles, en particulier lorsque la durée de suivi est longue. Avec peu de données manquantes et une collecte du statut vital des individus auprès de l'état civil de la commune de naissance du patient ou par recoupement avec le registre CépiDc, le SNDS est donc particulièrement adapté aux études nécessitant un long suivi. Enfin, les données de remboursement permettent de disposer d'éléments précis pour une évaluation économique. Néanmoins, les feuilles de soins peuvent être présentées au remboursement parfois plusieurs mois après la réalisation de l'acte.

Certaines données non collectées dans la base principale du SNDS sont toutefois des limites majeures à prendre en compte dans la réalisation des études. En particulier, l'absence de résultat d'examen clinique (imagerie, données biologiques...) ou paraclinique (tabagisme, tension, IMC...), l'absence de motifs de consultation et le manque de données sociales (hormis l'affiliation à la CMUC) limitent notamment l'identification de certains facteurs de risque. Le SNDS ne recense pas non plus les données de consommation des médicaments délivrés au cours des séjours hospitaliers (hors ceux inscrits sur la liste en sus). Les informations sur des diagnostics médicaux restent limitées au codage des séjours hospitaliers réalisé dans le cadre de la tarification à l'activité ou de l'inscription sur la liste des ALD, avec un risque de mauvaise classification si un patient est éligible à plusieurs ALD.

Les paragraphes suivants concernant l'utilisation d'algorithmes d'identification d'évènements de santé ou de maladies et les études multisources apportent quelques éléments de réponses aux limites évoquées ci-dessus.

Enfin, il convient de noter que la manipulation à des fins épidémiologiques du SNDS nécessite une connaissance et une compétence approfondies de ces bases de données.

Le recours à des algorithmes pour l'identification des pathologies

Le recours à des algorithmes d'identification est nécessaire par exemple pour définir un critère de jugement ou une variable jugée source de confusion potentielle, ou l'indication présumée du médicament ou du dispositif médical d'intérêt.

La direction de la Stratégie, des Études et des Statistiques – département des Études sur les Pathologies de la Caisse nationale de l'assurance maladie publie régulièrement des mises à jour d'algorithmes de repérage à partir des données de la base principale du SNDS pour des pathologies, évènements de santé ou traitements qu'elle utilise pour la cartographie des pathologies dans le cadre de son rapport annuel ou pour les fiches de « pathologie » [publiées sur le site ameli.fr](#). Leur méthodologie de construction peut être discutée mais la principale difficulté reste la validation de ces algorithmes qui nécessite un chaînage entre les données de la base principale du SNDS avec une autre source de données (cohorte, registre...) pour notamment en calculer la sensibilité et la spécificité.

Le réseau ReDSiam (réseau données Sniiram) a pour mission de développer, soumettre à une discussion critique puis mettre à disposition des algorithmes (84),(85).

Du fait de l'importance du problème, le *Health Data Hub* a lancé un appel à manifestation d'intérêt pour le développement d'algorithmes et leur validation.

Ces algorithmes ont aussi été la base du développement d'un indice mesurant la gravité de l'état de santé, avec l'intérêt d'une évaluation de leur capacité à prédire la mortalité chez les individus de plus de 65 ans affiliés au régime général (86).

Les études cliniques ou les registres sont une opportunité de tester les algorithmes mis à disposition, sous réserve d'un chaînage avec la base principale du SNDS (voir ci-dessous). Ainsi, par exemple, l'évaluation des performances d'un algorithme d'identification des hémorragies dans le contexte d'un

traitement par anticoagulant oral (87) a montré une faible sensibilité et une inhomogénéité en fonction de la nature de l'hémorragie (88).

Les algorithmes utilisés dans les études post-inscription devront être détaillés dans les protocoles soumis à la HAS.

L'intérêt des études multisources

Le chaînage entre la base principale du SNDS et d'autres sources de données a plusieurs intérêts (89). Tout d'abord, il permet de compléter les données de la base principale du SNDS par des données cliniques ou paracliniques afin notamment de prendre en compte les facteurs de risque dans les analyses statistiques. La cohorte CONSTANCES (90), l'étude CANARI (91), l'étude SACHA (92) ou encore l'étude ISO-PSY (93) sont des exemples intéressants d'études multisources. De plus, le chaînage de la base principale du SNDS à d'autres sources de données peut également permettre la validation d'algorithmes (94),(88), la réalisation à moindre coût d'un suivi à long terme, par exemple dans la suite d'un essai clinique randomisé, ou encore la collecte, là encore à moindre coût, des informations utiles pour une analyse économique des données issues d'un essai clinique randomisé.

Le chaînage reste néanmoins une entreprise complexe (95) malgré la simplification récente de la demande d'autorisation de l'utilisation du NIR par la Cnil (96). En l'absence de NIR, diverses modalités de chaînage sont proposées, indirectes (probabilistes ou semi-déterministes) (97). D'autres pistes font l'objet de recherche, sachant les contraintes spécifiques de l'exercice dans lequel les deux bases à chaîner peuvent être volumineuses et asymétriques en taille.

Par ailleurs, rendre plus simple le chaînage de bases de données avec la base principale du SNDS s'inscrit dans les missions et les objectifs du *Health Data Hub* avec le concours de la Caisse nationale de l'assurance maladie et de la Caisse nationale de l'assurance vieillesse.

Annexe 6. Mesures centrées sur le patient : l'exemple des *Patient Reported Outcome Measures*

Description et utilité

Le résultat perçu par le patient est un concept désigné en anglais par *Patient Reported Outcome* (PRO). Les instruments permettant de mesurer ces PROs (Patient Reported Outcome Measures – PROMs) sont majoritairement des auto-questionnaires permettant, en général, l'évaluation quantitative de nombreux concepts. Ces instruments mesurent une échelle unidimensionnelle ou multidimensionnelle. Chaque échelle est composée d'un ou de plusieurs items (des questions) disposant d'un format de réponse préspecifié. Pour une échelle donnée, la mesure de la dimension (par exemple le niveau de fatigue) – une unique valeur quantitative – est obtenue par une transformation algébrique des réponses aux items : le modèle de mesure.

Les PROMs se distinguent des autres types de résultats évaluables en santé par le fait que la mesure est réalisée par le patient, sans interprétation des réponses par un professionnel de santé (98). Ainsi, les PROMs sont un instrument de choix pour mesurer des concepts selon la perspective du patient (99), dont les concepts subjectifs (relatifs au sujet pensant, comme la douleur, la fatigue, l'anxiété, et plus largement la qualité de vie liée à la santé). S'il est validé de façon adéquate, un PROM dispose alors de bonnes propriétés métrologiques (psychométriques). Dans le cadre des études en vie réelle, ils sont une ressource utile pour enrichir l'évaluation de l'impact des technologies de santé selon la perspective des patients. Toutefois, l'utilisation de ces instruments est contingente de problématiques méthodologiques spécifiques.

Considérations méthodologiques

Choix du concept à mesurer

Fréquemment, les PROMs sont désignés comme « les échelles de qualité de vie », ce qui peut parfois être impropre. La qualité de vie (ou la qualité de vie liée à la santé) n'est qu'un des concepts mesurables par un PRO. Certains sont des évaluations de l'état de santé : par exemple d'intensité de symptômes, de déficience ou de fonctionnement. L'EQ-5D, un questionnaire largement utilisé, en est un exemple. D'autres PROMs sont relatifs à la satisfaction des patients par rapport aux soins reçus. D'autres encore abordent la santé perçue comme le ressenti subjectif de l'état de santé ou le fonctionnement social. Ces derniers instruments peuvent alors mesurer ce que l'on appelle la qualité de vie liée à la santé. De nombreux PROMs – dits spécifiques – sont développés pour une pathologie donnée, d'autres sont dits génériques. Un avantage des instruments génériques est de disposer d'une mesure permettant de comparer différentes populations ou situations pathologiques. Les instruments spécifiques disposent fréquemment d'une meilleure validité de contenu (ils couvrent de façon plus exhaustive la pathologie donnée) (100).

Finalement, il faut sélectionner rigoureusement un PROM au regard du concept que l'on souhaite mesurer. On peut trouver une taxonomie des concepts évaluables en santé via la classification ICF de l'OMS (101), le modèle biopsychosocial de Wilson et Cleary (102), ou bien le *Montreal Accord on Patient Reported Outcomes* (103). PROQOLID™ (<https://eprovide.mapi-trust.org/>), maintenue par *MAPI Research Trust*, est une base de données de recherche de PROMs. Le panorama de la HAS sur les PREMs et PROMs (104) ainsi que les guides de bonnes pratiques (105) peuvent également orienter le choix de l'auto-questionnaire. Des considérations pratiques doivent aussi guider la sélection de l'instrument comme la longueur de passation ou la charge cognitive associée à la passation (niveau de langage, format de réponse...) ou la conformité aux guides méthodologiques de la HAS.

Qualité de l'outil utilisé

Les PROMs ont le plus souvent comme objectif de proposer une mesure quantitative qui doit présenter de bonnes qualités psychométriques. Il existe deux grandes propriétés psychométriques :

1. la validité : l'instrument mesure-t-il ce qu'il est censé mesurer ? (par exemple, vérifier qu'un instrument mesurant de la symptomatologie dépressive ne capture pas de la symptomatologie anxieuse) ;
2. la fiabilité : que vaut la mesure ? (par exemple, vérifier que dans des conditions identiques, un instrument rend la même valeur).

La vérification des propriétés psychométriques d'un PROM est étudiée spécifiquement lors de sa création et validation, puis complétée par des travaux ultérieurs en utilisation courante (106). Les travaux spécifiques consistent, en plus de la création de l'instrument et d'une passation exploratoire sur quelques sujets (accompagnée d'études qualitatives), en une validation chez quelques centaines de sujets puis en une analyse statistique des propriétés psychométriques. Fiabilité et validité ne s'apprécient pas via un unique indicateur, mais par de multiples critères, en accumulant des preuves au fur et à mesure de l'utilisation de l'instrument. La validité d'un instrument comporte de nombreux sous-aspects : validité de face, de contenu, de construit, structurelle, de critère, etc. Le groupe international COSMIN¹⁰ propose une taxonomie consensuelle des propriétés psychométriques (107). Ne doivent être utilisés que des PROs qui ont fait preuve de propriétés psychométriques suffisantes. Une revue systématique de la bibliographie ainsi que l'analyse critique de preuves doivent être conduites avant utilisation d'un PROM.

Deux aspects spécifiques sont à noter. Il faut s'assurer que le PROM souhaité a été validé dans la langue d'utilisation. La traduction d'un PROM dans une autre langue que sa langue d'origine (l'adaptation transculturelle) suit des règles spécifiques, y compris une validation dans la langue de traduction, car c'est un processus à fort risque de dégrader les propriétés psychométriques par variabilité interculturelle (108) (<https://euroqol.org/publications/user-guides/>). Il faut s'assurer d'une adéquation entre la population cible de l'étude en vie réelle et la population de validation du PROM. Entre autres, une validation dans une population hospitalière ne garantit pas une adéquate utilisation en communauté. Il est préférable d'utiliser les données de l'étude en vie réelle pour révéifier les propriétés psychométriques de l'outil et confirmer une utilisation adéquate.

Utilisation pratique

Si le format papier reste prédominant, l'utilisation de PROMs sous format électronique (ePROs) est croissante (109). Outre l'intérêt logistique (absence de saisie manuelle des données), les ePROs présentent l'intérêt d'une plus facile administration dans des contextes variés (par exemple, à domicile via une interface électronique). Le dispositif de réponse peut être adapté à des déficiences (visuelle par exemple). Certains instruments électroniques, dits CAT (pour *Computerized Adaptive Testing*), permettent une adaptation des items proposés (en nombre et contenu) parmi une banque préconçue afin de mesurer le concept d'intérêt avec le moins d'items possible (110). Néanmoins, ces instruments CAT nécessitent l'utilisation d'un modèle de mesure complexe et adéquat, requérant de la modélisation des réponses aux items.

Il est primordial, pour obtenir une mesure valide et fiable du concept d'intérêt, d'utiliser un instrument rigoureusement selon son modèle de mesure proposé. Il existe plusieurs types de modèles de mesure. Le modèle de mesure le plus utilisé est la théorie classique des tests (*Classical Test Theory*, CTT). La mesure est souvent la simple somme des items d'une échelle (parfois pondérée). Cette mesure facilement obtenue est appelée le score et, sous réserve de bonnes propriétés psychométriques, est une

¹⁰ *CO*nsensus-based Standards for the selection of health Measurement *IN*struments.

bonne mesure ordinale (elle permet d'ordonner des sujets sur l'échelle) du concept d'intérêt (111). D'autres modèles de mesure sont dits à variable latente (une variable non observable mais dont on suppose qu'elle explique les réponses de plusieurs items, par exemple le concept de qualité de vie) : modèles de Rasch (*Rasch Measurement Theory*, RMT), modèles de réponses à l'item (*Item Response Theory*, IRT) ou modèles à équations structurelles (*Structural Equation Modeling*, SEM) (112, 113). Ces modèles, au prix d'une modélisation numérique, permettent l'obtention d'une mesure aux propriétés plus intéressantes comme la propriété d'intervalle (chaque écart d'une unité sur l'échelle représente une même quantité) ou bien la modélisation de structures complexes. De plus, pour une échelle donnée, il ne faut pas retirer d'items de l'échelle (exception faite des instruments CAT). Pour un PROM multidomaine, il est licite d'estimer une mesure globale en plus du profil de mesures de chacune des échelles uniquement si la structure de l'instrument le prévoit et qu'il existe un modèle de mesure de cette valeur.

Résultats et interprétation

L'interprétation des résultats issus des PROMs (notamment l'appréciation de la quantité d'effet), que ce soit une estimation ponctuelle, une différence intergroupe ou un changement, n'est parfois pas aisée. Certains PROMs ont été administrés dans de grands échantillons et on dispose ainsi de distributions que l'on peut qualifier de normes populationnelles. C'est le cas du SF-36, une mesure de qualité de vie, pour laquelle la distribution dans de nombreuses populations est disponible. Dans ce cas, une comparaison d'un sujet ou de groupes par rapport à la norme populationnelle peut aider à l'interprétation (114). C'est le cas aussi de l'instrument EQ-5D qui est un instrument multi-attribut à considérer pour l'évaluation économique (115, 116).

Par ailleurs, l'utilisation *a priori* d'une définition d'un répondeur (*Responder Definition*, RD) est fréquente. Il s'agit d'un seuil de changement sur une période prédéterminée qui peut être interprété comme un bénéfice du traitement. On classe ainsi chaque patient d'un échantillon comme ayant expérimenté un bénéfice ou non (98). Cette RD peut être obtenue selon de multiples perspectives : cliniques selon le patient (une valeur qui associe le changement sur l'échelle du PRO à un *outcome* comme la sévérité de la maladie ou du pronostic) ou un professionnel (le changement minimal qui sera considéré comme un bénéfice significatif). Une perspective fréquemment utilisée est d'évaluer la différence minimale pertinente (*Minimal Importance Difference*, MID) d'un PRO : soit le plus petit changement que le sujet perçoit comme significatif (117). Certains instruments sont ainsi accompagnés d'estimations de MID, ce qui peut être une aide à l'interprétation.

La réponse aux items étant le résultat d'un processus interprétatif variable, il est de plus en plus constaté que les PROMs ne résistent pas à l'hypothèse d'invariance de mesure (la stabilité de l'échelle de mesure, que ce soit au cours du temps ou entre populations différentes). Au cours du temps, du fait de mécanismes d'adaptation psychologique aux événements de santé, la représentation qu'ont les patients des concepts que l'on mesure peut changer : la conséquence est le phénomène de *response shift* (par exemple, être « modérément fatigué » ne correspond pas au même niveau de fatigue avant et après chimiothérapie (118)). Des différences systématiques de niveau de réponse à certains items d'un PRO peuvent exister entre différentes populations (par exemple, selon des catégories d'âge, ou selon le genre) à même niveau de concept. Pour un même niveau de fatigue, des sujets âgés peuvent répondre différemment à une question que des sujets jeunes, par interprétation différente de la question. Ceci est connu sous le nom de fonctionnement différentiel des items (*Differential item Functioning*, DIF (119)). Ces violations d'invariance peuvent invalider la comparaison simple de scores comme représentant correctement des variations du niveau du concept d'intérêt. Ces violations peuvent être prises en compte et enrichir l'interprétation des variations de mesure via l'utilisation de méthodes et de questionnaires appropriés.

Recherche documentaire

Une recherche documentaire a été conduite dans les bases de données Medline, Embase et Science Direct de janvier 2013 à avril 2020, en anglais et en français, à partir de l'équation de recherche suivante : *"Product Surveillance, Postmarketing"[Mesh] AND ("Drug Approval"[Majr] OR "Device Approval"[Majr]) OR (post-approval study or post-approval studies OR real world study design or real world studies design OR real life study design or real life studies design OR post-authorization studies or post-authorization study OR observational stud* design) Field: Title.*

281 références ont été obtenues.

Références bibliographiques

1. Haute Autorité de Santé. Doctrine de la commission de la transparence. Principes d'évaluation de la CT relatifs aux médicaments en vue de leur accès au remboursement. Saint-Denis La Plaine: HAS; 2018. https://www.has-sante.fr/upload/docs/application/pdf/2018-10/doctrine_10102018.pdf
2. Haute Autorité de Santé. Principes d'évaluation de la CNEDiMTS relatifs aux dispositifs médicaux à usage individuel en vue de leur accès au remboursement. Saint-Denis La Plaine: HAS; 2019. https://www.has-sante.fr/upload/docs/application/pdf/2017-11/principes_devaluation_de_la_cnedimts-v4-161117.pdf
3. Haute Autorité de Santé. Choix méthodologiques pour l'évaluation économique à la HAS. Saint-Denis La Plaine: HAS; 2019. https://www.has-sante.fr/upload/docs/application/pdf/2020-07/guide_methodologique_evaluation_economique_has_2020_vf.pdf
4. Haute Autorité de Santé. Choix méthodologiques pour l'analyse de l'impact budgétaire à la HAS. Saint-Denis La Plaine: HAS; 2016. https://www.has-sante.fr/upload/docs/application/pdf/2016-12/guide_methodologique_choix_methodologiques_pour_l_analyse_de_limpact_budgetaire_a_la_has_.pdf
5. Haute Autorité de Santé. Cartographie des impacts organisationnels pour l'évaluation des technologies de santé. Saint-Denis La Plaine: HAS; 2020. https://www.has-sante.fr/upload/docs/application/pdf/2020-12/guide_methodologique_impacts_organisationnels.pdf
6. Parlement Européen, Conseil de l'Union Européenne. Règlement (UE) 2016/679 du parlement européen et du conseil du 27 avril 2016 relatif à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données, et abrogeant la directive 95/46/CE (règlement général sur la protection des données) Journal Officiel de l'Union Européenne 2016; L 119-1(04/05).
7. Parlement Européen, Conseil de l'Union Européenne. Rectificatif au règlement (UE) 2016/679 du Parlement européen et du Conseil du 27 avril 2016 relatif à la protection des personnes physiques à l'égard du traitement des données à caractère personnel et à la libre circulation de ces données, et abrogeant la directive 95/46/CE (règlement général sur la protection des données). Journal Officiel de l'Union Européenne 2018 ;L 119-1(23/05).
8. Agence du numérique en santé. Hébergement des données de santé. Labels et certifications [En ligne] 2021. <https://esante.gouv.fr/labels-certifications/hebergement-des-donnees-de-sante>
9. Conseil national de l'Industrie, Levy-Bachelot L, Mollo L, Josserean A, Bacry E, Combes S, *et al.* Programme «intelligence Artificielle et santé» du Comité Stratégique de Filière Industries et Technologies de Santé. Paris: Alliance Recherche et Innovation Industries Santé (ARRIS); 2021. https://gt2.ariis.fr/wp-content/uploads/2021/03/Recommandations-du-GT-Usage-des-donnees-en-vie-reelle-dans-la-prise-de-decision_vf.pdf
10. Haute Autorité de Santé. Les études post-inscription sur les technologies de santé (médicaments, dispositifs médicaux et actes) Principes et méthodes. Saint-Denis La Plaine: HAS; 2011. https://www.has-sante.fr/upload/docs/application/pdf/2012-01/etudes_post_inscription_technologies_sante.pdf
11. Haute Autorité de Santé. Projet stratégique 2019 – 2024. Saint-Denis La Plaine: HAS; 2018. https://www.has-sante.fr/upload/docs/application/pdf/2018-11/projet_strategique_2019-2024.pdf
12. Makady A, Ham RT, de Boer A, Hillege H, Klungel O, Goettsch W. Policies for use of real-world data in health technology assessment (hta): a comparative study of six HTA Agencies. *Value Health* 2017;20(4):520-32. <http://dx.doi.org/10.1016/j.jval.2016.12.003>
13. Makady A, van Veelen A, Jonsson P, Moseley O, D'Andon A, de Boer A, *et al.* Using real-world Data in Health Technology Assessment (HTA) practice: A comparative study of five HTA agencies. *Pharmacoeconomics* 2018;36(3):359-68. <http://dx.doi.org/10.1007/s40273-017-0596-z>
14. Hordijk-Trion M, Lenzen M, Wijns W, de Jaegere P, Simoons ML, Scholte OP, Reimer WJ, *et al.* Patients enrolled in coronary intervention trials are not representative of patients in clinical practice: results from the Euro Heart Survey on Coronary Revascularization. *Eur Heart J* 2006;27(6):671-8. <http://dx.doi.org/10.1093/eurheartj/ehi731>
15. Lewis JH, Kilgore ML, Goldman DP, Trimble EL, Kaplan R, Montello MJ, *et al.* Participation of patients 65 years of age or older in cancer clinical trials. *J Clin Oncol* 2003;21(7):1383-9.
16. Van Spall HG, Toren A, Kiss A, Fowler RA. Eligibility criteria of randomized controlled trials published in high-impact general medical journals: a systematic sampling review. *JAMA* 2007;297(11):1233-40. <http://dx.doi.org/10.1001/jama.297.11.1233>
17. Haute Autorité de Santé. Sexe, genre et santé. Rapport d'analyse prospective 2020. Saint-Denis La Plaine: HAS; 2020. https://www.has-sante.fr/upload/docs/application/pdf/2020-12/rapport_analyse_prospective_2020.pdf
18. Gor KA, Shah KN, Joshi PB, Joshi HM, Rana DA, Malhotra SD. Off-label drugs use in neurology outpatient department: A prospective study at a tertiary care teaching hospital. *Perspect Clin Res* 2020;11(1):31-6. http://dx.doi.org/10.4103/picr.PICR_117_18
19. Rebordosa C, Houben E, Laugesen K, Bothner U, Montonen J, Aguado J, *et al.* No evidence of off-label use of olodaterol and indacaterol in Denmark, France, and the Netherlands: A Drug Utilization Study. *Sci Rep* 2020;10(1):586. <http://dx.doi.org/10.1038/s41598-019-57397-5>
20. Stafford RS. Regulating off-label drug use--rethinking the role of the FDA. *N Engl J Med* 2008;358(14):1427-9. <http://dx.doi.org/10.1056/NEJMp0802107>
21. Briasoulis A, Gao Y, Inampudi C, Alvarez P, Asleh R, Chrischilles E, *et al.* Characteristics and outcomes in

- patients with atrial fibrillation receiving direct oral anticoagulants in off-label doses. *BMC Cardiovasc Disord* 2020;20(1):42.
<http://dx.doi.org/10.1186/s12872-020-01340-4>
10.1186/s12872-020-01340-4.
22. Nordon C, Karcher H, Groenwold RH, Ankarfeldt MZ, Pichler F, Chevrou-Severac H, *et al*. The "efficacy-effectiveness gap": Historical background and current conceptualization. *Value Health* 2016;19(1):75-81.
<http://dx.doi.org/10.1016/j.jval.2015.09.2938>
23. Chassang S, Snowberg E, Seymour B, Bowles C. Accounting for behavior in treatment effects: new applications for blind trials. *PLoS ONE* 2015;10(6):e0127227.
<http://dx.doi.org/10.1371/journal.pone.0127227>
10.1371/journal.pone.0127227. eCollection 2015.
24. Nordon C, Bovagnet T, Belger M, Jimenez J, Olivares R, Chevrou-Severac H, *et al*. Trial exclusion criteria and their impact on the estimation of antipsychotic drugs effect: A case study using the SOHO database. *Schizophr Res* 2018;193:146-53.
<http://dx.doi.org/10.1016/j.schres.2017.07.031>
10.1016/j.schres.2017.07.031. Epub 2017 Jul 13.
25. Rothwell PM. Factors that can affect the external validity of randomised controlled trials. *PLoS Clin Trials* 2006;1(1):e9.
<http://dx.doi.org/10.1371/journal.pctr.0010009>
26. Kennedy-Martin T, Curtis S, Faries D, Robinson S, Johnston J. A literature review on the representativeness of randomized controlled trial samples and implications for the external validity of trial results. *Trials* 2015;16:495.
<http://dx.doi.org/10.1186/s13063-015-1023-4>
10.1186/s13063-015-1023-4.
27. Haute Autorité de Santé, Commission de la Transparence. Doctrine de la Commission de la Transparence. Principes d'évaluation de la CT relatifs aux médicaments en vue de leur accès au remboursement. Saint-Denis La Plaine: HAS; 2020.
https://www.has-sante.fr/upload/docs/application/pdf/2021-03/doctrine_ct.pdf
28. Fleming TR, Powers JH. Biomarkers and surrogate endpoints in clinical trials. *Stat Med* 2012;31(25):2973-84.
<http://dx.doi.org/10.1002/sim.5403>
29. European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP). Code of conduct. London: ENCEPP; 2018.
http://www.encepp.eu/standards_and_guidances/checkListProtocols.shtml
30. European Network for Health Technology Assessment, Haute Aurorité de Santé. Criteria to select and prioritize health technologies for additional evidence generation. London: EUnetHTA; 2012.
31. European Network for Health Technology Assessment, Haute Aurorité de Santé, Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Position paper on how to best formulate research recommendations for primary research arising from HTA reports. London: EUnetHTA; 2015.
32. European Network for Health Technology Assessment, Haute Aurorité de Santé, Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen. Position paper on how to decide on the appropriate study design for primary research arising from HTA reports. London: EUnetHTA; 2015.
33. European Network for Health Technology Assessment. Core protocol pilot for additional evidence generation. London: EUnetHTA; 2019.
https://eunetha.eu/wp-content/uploads/2019/10/EUnetHTA_Core-protocol-Pilot-for-AEG_0.pdf
34. European Network for Health Technology Assessment. REQueST® Tool and its vision paper [En ligne]. London: EUnetHTA; 2020.
<https://eunetha.eu/wp-content/uploads/2019/10/EUnetHTA-Position-Paper-on-research-recommendations.pdf>
35. European Network for Health Technology Assessment. EUnetHTA PLEG activities [En ligne] 2018.
<https://eunetha.eu/pleg/>
36. European Medicines Agency, Arlett P. DARWIN EU (Data Analytics and Real World Interrogation Network). PCWP and HCPWP Data workshop 23 September 2020. Amsterdam: EMA; 2020.
https://www.ema.europa.eu/en/documents/presentation/presentation-proposal-darwin-eu-data-analytics-real-world-interrogation-network-parlett-ema_en.pdf
37. European Medicines Agency. Guideline on registry-based studies - launch of public consultation [En ligne] 2020.
<https://www.ema.europa.eu/en/news/guideline-registry-based-studies-launch-public-consultation>
38. International Society for Pharmacoepidemiology. Guidelines for Good Pharmacoepidemiology Practices (GPP). Bethesda: ISPE; 2015.
39. von Elm E, Altman DG, Egger M, Pocock SJ, Gotsche PC, Vandenbroucke JP, *et al*. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *J Clin Epidemiol* 2008;61(4):344-9.
<http://dx.doi.org/10.1016/j.jclinepi.2007.11.008>
10.1016/j.jclinepi.2007.11.008.
40. Gallo V, Egger M, McCormack V, Farmer PB, Ioannidis JP, Kirsch-Volders M, *et al*. Strengthening the Reporting of OBservational studies in Epidemiology--Molecular Epidemiology (STROBE-ME): an extension of the STROBE Statement. *PLoS Med* 2011;8(10):e1001117.
<http://dx.doi.org/10.1371/journal.pmed.1001117>
41. The European Network of Centres for Pharmacoepidemiology and Pharmacovigilance. Guide on methodological standards in Pharmacoepidemiology (revision 8) : ENEPP; 2010.
42. European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP). ENCePP Checklist for study protocols (Revision 4). London: ENCEPP; 2018.
http://www.encepp.eu/standards_and_guidances/checkListProtocols.shtml
43. Berger ML, Sox H, Willke RJ, Brixner DL, Eichler HG, Goettsch W, *et al*. Good practices for real-world data studies of treatment and/or comparative effectiveness: recommendations from the joint ISPOR-ISPE special task force on real-world evidence in health care decision making. *Value Health* 2017;20(8):1003-8.
<http://dx.doi.org/10.1016/j.jval.2017.08.3019>
10.1016/j.jval.2017.08.3019. Epub 2017 Sep 15.

44. Agency for Healthcare Research and Quality, Velentgas P, Dreyer NA, Nourjah P, Smith SR, Torchia MM. Developing a protocol for observational comparative effectiveness research. A user's guide. Rockville: AHRQ; 2013.
<https://effectivehealthcare.ahrq.gov/sites/default/files/related-files/user-guide-observational-cer-130113.pdf>
45. Schulz KF, Altman DG, Moher D, Group C. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMJ* 2010;340:c332.
<http://dx.doi.org/10.1136/bmj.c332>
10.1136/bmj.c332.
46. Benchimol EI, Smeeth L, Guttman A, Harron K, Moher D, Petersen I, *et al.* The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) statement. *PLoS Med* 2015;12(10):e1001885.
<http://dx.doi.org/10.1371/journal.pmed.1001885>
47. Wang SV, Schneeweiss S, Berger ML, Brown J, de Vries F, Douglas I, *et al.* Reporting to improve reproducibility and facilitate validity assessment for healthcare database studies V1.0. *Pharmacoepidemiol Drug Saf* 2017;26(9):1018-32.
<http://dx.doi.org/10.1002/pds.4295>
48. Berdai D, Thomas-Delecout F, Szwarcensztein K, d'Andon A, Collignon C, Comet D, *et al.* Requests for post-registration studies (PRS), patients follow-up in actual practice: Changes in the role of databases. *Therapie* 2018;73(1):13-24.
<http://dx.doi.org/10.1016/j.therap.2017.12.008>
49. Food and Drug Administration, Center for Devices and Radiological Health, Center for Biologics Evaluation and Research. Use of real-world evidence to support regulatory decision-making for medical devices. Guidance for Industry and Food and Drug Administration Staff. Rockville: FDA; 2017.
<https://www.fda.gov/media/99447/download>
50. European Network for Health Technology Assessment. Therapeutic medical devices. Guideline. London: EUnetHTA; 2015.
https://eunetha.eu/wp-content/uploads/2018/01/Therapeutic-medical-devices_Guideline_Final-Nov-2015.pdf
51. Shea BJ, Reeves BC, Wells G, Thuku M, Hamel C, Moran J, *et al.* AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *BMJ* 2017;358:j4008.
<http://dx.doi.org/10.1136/bmj.j4008>
52. Stroup DF, Berlin JA, Morton SC, Olkin I, Williamson GD, Rennie D, *et al.* Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. *JAMA* 2000;283(15):2008-12.
<http://dx.doi.org/10.1001/jama.283.15.2008>
53. Wang SV, Pinheiro S, Hua W, Arlett P, Uyama Y, Berlin JA, *et al.* STaRT-RWE: structured template for planning and reporting on the implementation of real world evidence studies. *BMJ* 2021;372:m4856.
<http://dx.doi.org/10.1136/bmj.m4856>
54. Carroll R, Ramagopalan SV, Cid-Ruzafa J, Lambrelli D, McDonald L. An analysis of characteristics of post-authorisation studies registered on the ENCePP EU PAS Register. *F1000Res* 2017;6:1447.
<http://dx.doi.org/10.12688/f1000research.12198.2>
10.12688/f1000research.12198.2. eCollection 2017.
55. Hill KP, Ross JS, Egilman DS, Krumholz HM. The ADVANTAGE seeding trial: a review of internal documents. *Ann Intern Med* 2008;149(4):251-8.
<http://dx.doi.org/10.7326/0003-4819-149-4-200808190-00006>
10.7326/0003-4819-149-4-200808190-00006.
56. Berger ML, Sox H, Willke RJ, Brixner DL, Eichler HG, Goettsch W, *et al.* Good practices for real-world data studies of treatment and/or comparative effectiveness: Recommendations from the joint ISPOR-ISPE Special Task Force on real-world evidence in health care decision making. *Pharmacoepidemiol Drug Saf* 2017;26(9):1033-9.
<http://dx.doi.org/10.1002/pds.4297>
57. Strom BL, Kimmel SE, Hennessy S. *Pharmacoepidemiology*. 6th ed. : Wiley-Blackwell; 2019.
58. Rothman KJ, Greenland S, Lash TL. *Modern Epidemiology*. 3rd ed. Philadelphia: Lippincott Williams and Wilkins; 2012.
59. Lund JL, Richardson DB, Sturmer T. The active comparator, new user study design in pharmacoepidemiology: historical foundations and contemporary application. *Curr Epidemiol Rep* 2015;2(4):221-8.
<http://dx.doi.org/10.1007/s40471-015-0053-5>
10.1007/s40471-015-0053-5. Epub 2015 Sep 30.
60. Levesque LE, Hanley JA, Kezouh A, Suissa S. Problem of immortal time bias in cohort studies: example using statins for preventing progression of diabetes. *BMJ* 2010;340:b5087.
<http://dx.doi.org/10.1136/bmj.b5087>
10.1136/bmj.b5087.
61. Dickerman BA, García-Albéniz X, Logan RW, Denaxas S, Hernán MA. Avoidable flaws in observational analyses: an application to statins and cancer. *Nat Med* 2019;25(10):1601-6.
<http://dx.doi.org/10.1038/s41591-019-0597-x>
62. Sterne JA, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, *et al.* Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ* 2009;338:b2393.
<http://dx.doi.org/10.1136/bmj.b2393>
63. Concato J, Lawler EV, Lew RA, Gaziano JM, Aslan M, Huang GD. Observational methods in comparative effectiveness research. *Am J Med* 2010;123(12 Suppl 1):e16-23.
<http://dx.doi.org/10.1016/j.amjmed.2010.10.004>
64. Candlish J, Pate A, Sperrin M, van Staa T. Evaluation of biases present in the cohort multiple randomised controlled trial design: a simulation study. *BMC Med Res Methodol* 2017;17(1):17.
<http://dx.doi.org/10.1186/s12874-017-0295-7>
65. Pate A, Candlish J, Sperrin M, Van Staa TP. Cohort Multiple Randomised controlled trials (cmRCT) design: efficient but biased? A simulation study to evaluate the feasibility of the Cluster cmRCT design. *BMC Med Res Methodol* 2016;16(1):109.
<http://dx.doi.org/10.1186/s12874-016-0208-1>
66. Eichner FA, Groenwold RHH, Grobbee DE, Oude Rengerink K. Systematic review showed that stepped-wedge cluster randomized trials often did not reach their planned sample size. *J Clin Epidemiol* 2019;107:89-100.
<http://dx.doi.org/10.1016/j.jclinepi.2018.11.013>

67. Reams BD, O'Malley CD, Critchlow CW, Lauffenburger JC, Brookhart MA. Changing patterns of use of osteoporosis medications in the years after launch: implications for comparative effectiveness research. *Pharmacoepidemiol Drug Saf* 2014;23(3):251-60.
<http://dx.doi.org/10.1002/pds.3545>
10.1002/pds.3545. Epub 2013 Nov 25.
68. Ankarfeldt MZ, Thorsted BL, Groenwold RH, Adalsteinsson E, Ali MS, Klungel OH. Assessment of channeling bias among initiators of glucose-lowering drugs: A UK cohort study. *Clin Epidemiol* 2017;9:19-30.
<http://dx.doi.org/10.2147/CLEP.S124054>
69. Schneeweiss S, Rassen JA, Glynn RJ, Avorn J, Mogun H, Brookhart MA. High-dimensional propensity score adjustment in studies of treatment effects using health care claims data. *Epidemiology* 2009;20(4):512-22.
<http://dx.doi.org/10.1097/EDE.0b013e3181a663cc>
10.1097/EDE.0b013e3181a663cc.
70. Robins JM, Hernan MA, Brumback B. Marginal structural models and causal inference in epidemiology. *Epidemiology* 2000;11(5):550-60.
<http://dx.doi.org/10.1097/00001648-200009000-00011>
10.1097/00001648-200009000-00011.
71. Ford I, Norrie J. Pragmatic Trials. *N Engl J Med* 2016;375(5):454-63.
<http://dx.doi.org/10.1056/NEJMra1510059>
10.1056/NEJMra1510059.
72. Schwartz D, Lellouch J. Explanatory and pragmatic attitudes in therapeutical trials. *J Chronic Dis* 1967;20(8):637-48.
[http://dx.doi.org/10.1016/0021-9681\(67\)90041-0](http://dx.doi.org/10.1016/0021-9681(67)90041-0)
73. Thorpe KE, Zwarenstein M, Oxman AD, Treweek S, Furburg CD, Altman DG, *et al*. A pragmatic-explanatory continuum indicator summary (PRECIS): a tool to help trial designers. *J Clin Epidemiol* 2009;62(5):464-75.
<http://dx.doi.org/10.1016/j.jclinepi.2008.12.011>
74. Loudon K, Treweek S, Sullivan F, Donnan P, Thorpe KE, Zwarenstein M. The PRECIS-2 tool: designing trials that are fit for purpose. *BMJ* 2015;350:h2147.
<http://dx.doi.org/10.1136/bmj.h2147>
75. Staa TP, Goldacre B, Gulliford M, Cassell J, Pirmohamed M, Taweel A, *et al*. Pragmatic randomised trials using routine electronic health records: putting them to the test. *BMJ* 2012;344:e55.
<http://dx.doi.org/10.1136/bmj.e55>
10.1136/bmj.e55.
76. Giraudeau B, Ravaud P. Preventing bias in cluster randomised trials. *PLoS Med* 2009;6(5):e1000065.
<http://dx.doi.org/10.1371/journal.pmed.1000065>
10.1371/journal.pmed.1000065. Epub 2009 May 5.
77. Califf RM. Pragmatic clinical trials: Emerging challenges and new roles for statisticians. *Clinical trials (London, England)* 2016;13(5):471-7.
<http://dx.doi.org/10.1177/1740774516656944>
78. Cook AJ, DeLong E, Murray DM, Vollmer WM, Heagerty PJ. Statistical lessons learned for designing cluster randomized pragmatic clinical trials from the NIH Health Care Systems Collaboratory Biostatistics and Design Core. *Clinical trials (London, England)* 2016;13(5):504-12.
<http://dx.doi.org/10.1177/1740774516646578>
79. Ezzati A, Lipton RB. Machine Learning Predictive models can improve efficacy of clinical trials for alzheimer's disease. *J Alzheimer Dis*. 2020;74(1):55-63.
<http://dx.doi.org/10.3233/jad-190822>
80. Adam G, Rampásek L, Safikhani Z, Smirnov P, Haibe-Kains B, Goldenberg A. Machine learning approaches to drug response prediction: challenges and recent progress. *NPJ Precision Oncology* 2020;4:19.
<http://dx.doi.org/10.1038/s41698-020-0122-1>
81. Direction générale de l'offre de soins. H07 : Les registres à caractère épidémiologique. Paris: Ministère des solidarités et de la santé; 2019.
https://solidarites-sante.gouv.fr/IMG/pdf/dgso_h07_fiche_mig_registres_2019.pdf
82. Schneeweiss S, Avorn J. A review of uses of health care utilization databases for epidemiologic research on therapeutics. *J Clin Epidemiol* 2005;58(4):323-37.
<http://dx.doi.org/10.1016/j.jclinepi.2004.10.012>
10.1016/j.jclinepi.2004.10.012.
83. Tuppin P, Rudant J, Constantinou P, Gastaldi-Ménager C, Rachas A, de Roquefeuil L, *et al*. Value of a national administrative database to guide public decisions: From the système national d'information interrégimes de l'Assurance Maladie (SNIIRAM) to the système national des données de santé (SNDS) in France. *Rev Epidemiol Sante Publique* 2017;65 Suppl 4:S149-S67.
<http://dx.doi.org/10.1016/j.respe.2017.05.004>
84. Fosse-Edorh S, Rigou A, Morin S, Fezeu L, Mandereau-Bruno L, Fagot-Campagna A. Algorithmes basés sur les données médico-administratives dans le champ des maladies endocriniennes, nutritionnelles et métaboliques, et en particulier du diabète. *Rev Epidemiol Sante Publique* 2017;65 Suppl 4:S168-S73.
<http://dx.doi.org/10.1016/j.respe.2017.05.001>
85. Goldberg M, Carton M, Doussin A, Fagot-Campagna A, Heyndrickx E, Lemaître M, *et al*. Le réseau REDSIAM (Réseau données Sniiram) – Spécial REDSIAM. *Rev Epidemiol Sante Publique* 2017;65 Suppl 4:S144-S8.
<http://dx.doi.org/10.1016/j.respe.2017.06.001>
86. Constantinou P, Tuppin P, Fagot-Campagna A, Gastaldi-Ménager C, Schellevis FG, Pelletier-Fleury N. Two morbidity indices developed in a nationwide population permitted performing outcome-specific severity adjustment. *J Clin Epidemiol* 2018;103:60-70.
<http://dx.doi.org/10.1016/j.jclinepi.2018.07.003>
87. Maura G, Blotière PO, Bouillon K, Billionnet C, Ricordeau P, Alla F, *et al*. Comparison of the short-term risk of bleeding and arterial thromboembolic events in nonvalvular atrial fibrillation patients newly treated with dabigatran or rivaroxaban versus vitamin K antagonists: a French nationwide propensity-matched cohort study. *Circulation* 2015;132(13):1252-60.
<http://dx.doi.org/10.1161/circulationaha.115.015710>
88. Oger E, Botrel MA, Juchault C, Bouget J. Sensitivity and specificity of an algorithm based on medico-administrative data to identify hospitalized patients with major bleeding presenting to an emergency department. *BMC Med Res Methodol* 2019;19(1):194.
<http://dx.doi.org/10.1186/s12874-019-0841-6>
89. Scailteux LM, Droitcourt C, Balusson F, Nowak E, Kerbrat S, Dupuy A, *et al*. French administrative health care database (SNDS): The value of its enrichment. *Thérapie* 2019;74(2):215-23.

<http://dx.doi.org/10.1016/j.therap.2018.09.072>

90. Goldberg M, Carton M, Descatha A, Leclerc A, Roquelaure Y, Santin G, *et al.* CONSTANCES: a general prospective population-based cohort for occupational and environmental epidemiology: cohort profile. *Occup Environ Med* 2017;74(1):66-71.

<http://dx.doi.org/10.1136/oemed-2016-103678>

91. Scailteux LM, Balusson F, Vincendeau S, Rioux-Leclercq N, Nowak E. Rationale and design of the CANARI study: a case-control study investigating the association between prostate cancer and 5-alpha-reductase inhibitors for symptomatic benign prostate hypertrophy by linking SNIIRAM and pathology laboratories in a specific region in France. *Fundam Clin Pharmacol* 2018;32(1):120-9.

<http://dx.doi.org/10.1111/fcp.12320>

92. Bouget J, Balusson F, Scailteux LM, Maignan M, Roy PM, L'Her E, *et al.* Major bleeding with antithrombotic agents: a 2012-2015 study using the French nationwide Health Insurance database linked to emergency department records within five areas - rationale and design of SACHA study. *Fundam Clin Pharmacol* 2019;33(4):443-62.

<http://dx.doi.org/10.1111/fcp.12444>

93. Agence nationale de sécurité du médicament et des produits de santé. Rapport scientifique de fin d'étude Etude «ISOPSY2». Saint-Denis : ANSM; 2020.

https://www.ansm.sante.fr/var/ansm_site/storage/original/application/bd8e92bf4eff92e1c0543708dcdddbbc.pdf

94. Fuentes S, Cosson E, Mandereau-Bruno L, Fagot-Campagna A, Bernillon P, Goldberg M, *et al.* Identifying diabetes cases in health administrative databases: a validation study based on a large French cohort. *Int J Public Health* 2019;64(3):441-50.

<http://dx.doi.org/10.1007/s00038-018-1186-3>

10.1007/s00038-018-1186-3. Epub 2018 Dec 4.

95. Pratt NL, Mack CD, Meyer AM, Davis KJ, Hammill BG, Hamp C, *et al.* Data linkage in pharmacoepidemiology: A call for rigorous evaluation and reporting. *Pharmacoepidemiol Drug Safety* 2020;29(1):9-17.

<http://dx.doi.org/10.1002/pds.4924>

96. Commission nationale informatique et libertés. Modalités de circulation du NIR pour la recherche en sante aux fins d'appariement de données avec le SNDS. Guide pratique. Paris: CNIL; 2020.

https://www.cnil.fr/sites/default/files/atoms/files/guide_pratique_circuits_nir_recherche_en_sante.pdf

97. Bounebache SK, Quantin C, Benzenine E, Obozinski G, Rey G. Revue bibliographique des méthodes de couplage des bases de données : Applications et perspectives dans le cas des données de santé publique *J Soc Fr Stat* 2018;159 :79-123.

98. Food and Drug Administration, Center for Drug Evaluation and Research (CDER), Center for Biologics Evaluation and Research (CBER), Center for Devices and Radiological Health (CDRH). Guidance for industry patient-reported outcome measures: use in medical product development to support labeling claims. Silver Springs: U.S. Department of Health and Human Services; 2009.

<https://www.fda.gov/media/77832/download>

99. Ziegelstein RC. Personomics. *JAMA internal medicine* 2015;175(6):888-9.

<http://dx.doi.org/10.1001/jamainternmed.2015.0861>

100. Fayers PM, Machin D. Quality of life: The assessment, analysis and interpretation of patient-reported outcomes. 2nd ed. Chichester: Wiley; 2013.

101. World Health Organization. How to use the ICF. A practical manual for using the International Classification of Functioning, Disability and Health (ICF). Exposure draft for comment. Geneva: WHO; 2013.

<https://www.who.int/classifications/drafticfpracticalmanual.pdf>

102. Wilson IB, Cleary PD. Linking clinical variables with health-related quality of life. A conceptual model of patient outcomes. *JAMA* 1995;273(1):59-65.

103. Mayo NE, Figueiredo S, Ahmed S, Bartlett SJ. Montreal Accord on Patient-Reported Outcomes (PROs) use series - Paper 2: terminology proposed to measure what matters in health. *J Clin Epidemiol* 2017;89:119-24.

<http://dx.doi.org/10.1016/j.jclinepi.2017.04.013>

104. Haute Autorité de Santé. Indicateurs de qualité perçue par le patient de type PROMS et PREMS – Panorama d'expériences étrangères et principaux enseignements. Note de cadrage. Saint-Denis La Plaine: HAS; 2020.

https://www.has-sante.fr/upload/docs/application/pdf/2020-05/panorama-proms-prems_note_de_cadrage.pdf

105. Haute Autorité de Santé. Guide du parcours de soins. Bronchopneumopathie chronique obstructive. Saint-Denis La Plaine: HAS; 2019.

https://www.has-sante.fr/upload/docs/application/pdf/2020-01/app_323_guide_bpco_actu_2019_vf.pdf

106. Falissard B. Mesurer la subjectivité en santé. Perspective méthodologique et statistique. Paris: Elsevier Masson; 2008.

https://www.elsevier-masson.fr/mesurer-la-subjectivite-en-sante-9782294703171.html?gclid=EAlaIqobChMlob3l9OaT7wIVCJ7VCh3sCw99EAAYASAAEglijyFD_BwE&gclid=aw_ds

107. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, *et al.* The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol* 2010;63(7):737-45.

<http://dx.doi.org/10.1016/j.jclinepi.2010.02.006>

108. Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine* 2000;25(24):3186-91.

<http://dx.doi.org/10.1097/00007632-200012150-00014>

109. Egeter J, Hüfner K, Sztankay M, Holzner B, Sperner-Unterwöger B. Implementation of an electronic routine outcome monitoring at an inpatient unit for psychosomatic medicine. *J Psychosom Res* 2018;105:64-71.

<http://dx.doi.org/10.1016/j.jpsychores.2017.12.009>

110. Rose M, Bjorner JB, Fischer F, Anatchkova M, Gandek B, Klapp BF, *et al.* Computerized adaptive testing--ready for ambulatory monitoring? *Psychosom Med* 2012;74(4):338-48.

<http://dx.doi.org/10.1097/PSY.0b013e3182547392>

111. De Vet HC, Terwee CB, Mokkink LB, Knol DL. Measurement in medicine. A practical guide. Cambridge: Cambridge University Press; 2011.

https://assets.cambridge.org/9780521118200/frontmatter/9780521118200_frontmatter.pdf

112. de Ayala RJ. The theory and practice of item response theory. New York: The Guilford Press; 2008.

113. Brown TA. Confirmatory factor analysis for applied research. 2nd ed. New York: The Guilford Press; 2015.

114. Leplège A, Ecosse E, Pouchot J, Coste J, Perneger T. Le questionnaire MOS SF-36 : manuel de l'utilisateur et guide d'interprétation des scores. Paris: Editions ESTEM; 2001.

115. EuroQol Group. EuroQol--a new facility for the measurement of health-related quality of life. Health Policy 1990;16(3):199-208.
[http://dx.doi.org/10.1016/0168-8510\(90\)90421-9](http://dx.doi.org/10.1016/0168-8510(90)90421-9)

116. Szende A, Janssen B, Cabases J. Self-Reported Population Health: An international perspective based on EQ-5D. Dordrecht NL: Springer; 2014.
https://www.ncbi.nlm.nih.gov/books/NBK500356/pdf/Bookshelf_NBK500356.pdf

117. Revicki D, Hays RD, Cella D, Sloan J. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. J Clin Epidemiol 2008;61(2):102-9.
<http://dx.doi.org/10.1016/j.jclinepi.2007.03.012>

118. Sprangers MAG, Sajobi T, Vanier A, Mayo NE, Sawatzky R, Lix LM, *et al.* Response shift in results of patient-reported outcome measures: a commentary to the response shift-in sync working group initiative. Quality Life Res.2021.
<http://dx.doi.org/10.1007/s11136-020-02747-4>

119. Teresi JA, Fleishman JA. Differential item functioning and health assessment. Qual Life Res. 2007;16 Suppl 1:33-42.
<http://dx.doi.org/10.1007/s11136-007-9184-6>

Abréviations et acronymes

AMM	Autorisation de mise sur le marché
CEESP	Commission d'évaluation économique et de santé publique
CNEDiMITS	Commission nationale d'évaluation des dispositifs médicaux et des technologies de santé
CT	Commission de la transparence
DM	Dispositif médical
ENCePP	<i>European Network of Centres for Pharmacoepidemiology and Pharmacovigilance</i>
EUneHTA	<i>European Network for Health Technology Assessment</i>
HAS	Haute Autorité de santé
ISPE	Internal Society of Pharmcoepidemiology
MAR	<i>Missing At Random</i>
NMAR	<i>Not Missing At Random</i>
PAES	<i>Post-authorisation efficacy studies</i>
PASS	<i>Post-authorisation safety studies</i>
PROM	<i>Patient Reported Outcome Measure</i>
REQUEST	<i>REQueST Tool and its vision paper</i>
SNDS	Système national des données de santé

Retrouvez tous nos travaux sur
www.has-sante.fr

