

Collection dirigée
par Dr A. Charon et Dr N. Meton

université
PARIS-SACLAY

FACULTÉ DE
MÉDECINE



EDN

LA

MARTINGALE*

2^e édition

LCA

LECTURE CRITIQUE D'ARTICLE

Méthodes de biomédecine quantitative
pour réussir l'épreuve de LCA



Cours de méthodologie

Fiches flash

Questions isolées et dossiers de LCA

avec correction détaillée

Coordination
d'**ALEXANDRA ROUQUETTE**
et **JOSIANE WARSZAWSKI**



INTRODUCTION

Ce qu'il faut comprendre, ce qu'il faut apprendre, ce qu'il faut savoir calculer

Objectifs de cet ouvrage

La biomédecine quantitative est l'ensemble des méthodes qui permettent de comprendre les bases quantitatives du **raisonnement médical**, pour la **prévention**, le **diagnostic**, la **prise en charge thérapeutique**, et l'évaluation du **pronostic** du patient.

Les progrès de la médecine sont largement issus d'observations et d'interventions empiriques. La médecine moderne s'appuie sur les méthodes de biomédecine quantitative pour évaluer les pratiques préventives et thérapeutiques, en apportant le **niveau de preuve** le plus élevé possible pour les justifier.

L'enseignement couvert par cet ouvrage vise à apporter une initiation aux méthodes qui permettent d'acquérir des connaissances sur la **fréquence**, les **facteurs de risque**, les **causes**, les **mécanismes physiopathologiques**, l'**histoire naturelle** et le **pronostic** des maladies, et d'évaluer l'**efficacité**, la **tolérance** et la mise en œuvre des **stratégies préventives, diagnostiques et thérapeutiques**.

Cet enseignement a trois objectifs :

- **Un objectif général** : être capable d'exercer un **esprit critique sur la qualité** des sources d'information qui sont à l'origine des connaissances actuelles et en **évaluer le niveau de preuve**. **Cet esprit critique s'exerce à chaque instant de la pratique médicale**. Le médecin est en effet continuellement amené à faire des choix entre des pratiques anciennes dont le niveau de preuve n'a pas toujours été évalué de manière rigoureuse et des innovations diagnostiques ou thérapeutiques dont le recul n'est pas toujours suffisant pour en identifier toutes les retombées.
- **Un objectif opérationnel à moyen terme** : apporter les outils nécessaires pour réussir l'épreuve de **lecture critique d'article (LCA) de l'Examen Dématérialisé National (EDN)** ; ces bases sont indispensables pour réussir l'UE 12 de LCA (formation à la recherche) en DFASM 1, 2 et 3.
- **Un objectif opérationnel à plus long terme** : apporter une **initiation à la recherche clinique et épidémiologique** qui fera partie des activités professionnelles des étudiants qui se dirigeront après l'internat vers une carrière alliant la recherche à leur pratique de la médecine en ville ou au sein d'une structure hospitalière.

Cet ouvrage est découpé en trois grandes parties :

- **Partie I – METHODES** : dans cette partie sont enseignées toutes les connaissances à acquérir au cours de l'UE Biomédecine Quantitative en DFGSM3 car elles sont essentielles à la compréhension des notions nécessaires à la réussite de l'épreuve de LCA à l'EDN. Pour chaque chapitre, une **fiche synthétique** des points à connaître est associée.
- **Partie II – TYPOLOGIE DES ETUDES SELON LEURS OBJECTIFS** : dans cette partie sont détaillées les particularités méthodologiques des différents types d'études à connaître car susceptibles de tomber à l'épreuve de LCA de l'EDN. De manière identique, pour chaque chapitre, une **fiche synthétique** des points à connaître est associée.
- **Partie III – LECTURE CRITIQUE D'ARTICLE** : dans cette partie, trois chapitres spécifiques de la LCA traitent de la validité, de la pertinence, du niveau de preuve des résultats d'une étude.

Depuis la 2ème édition, à chaque chapitre sont associées des **exercices d'entraînement** sous forme de **questions individuelles** ou de **dossier de LCA** au format **EDN**.

Cet ouvrage couvre une grande partie des connaissances de deux items du programme revisité du 2^{ème} cycle des études médicales (R2C) :

■ **Item 20 : La méthodologie de la recherche en santé**

Intitulé connaissance	Rang	Chapitre
1. Définition et organisation de la recherche en santé	A	I-A
2. La méthodologie de la recherche en santé	A	I-A
3. La réglementation de la recherche en santé	A	I-H
4. Justification et problématique de recherche	B	I-A
5. Hypothèses de recherche	B	I-A
6. Méthodes adaptées à un objectif de recherche	B	I-A
7. Modalités d'échantillonnage	B	I-B
8. Recueil des données	B	I-B
9. Pertinence des données recueillies	B	I-B
10. Probabilités, variables aléatoires, lois de probabilités	B	I-C
11. Analyses descriptives : savoir calculer les paramètres statistiques	B	I-C
12. Tests statistiques : principes des tests d'hypothèse	B	I-C
13. Données de survie	B	I-C, I-D
14. Connaître la notion et le rôle de l'ajustement. Savoir interpréter et lire un résultat ajusté	B	I-C, I-E
15. Mettre en œuvre les bonnes pratiques d'usage de l'informatique de santé pour aider la recherche clinique et en particulier les essais cliniques	B	I-H
16. Facteur de risque, facteur étiologique, facteur pronostique, déterminants de santé, facteur exposition	A	I-A, II-B
17. Population cible et source, échantillon	A	I-A
18. Principaux types d'études observationnelles.	A	I-B
19. Etudes au schéma moins classique	B	I-B
20. Mesure de fréquence, Prévalence, Incidence (taux d'), Risque	A	I-D
21. Mesure de fréquence standardisée	B	I-D, I-G
22. Mesures d'association brutes ou ajustées, différence de moyennes, Réduction absolue de risque (RAR), réduction relative de risque (RRR), Nombre de sujet à traiter (NST) et Rapport de taux d'incidence brut.	A	I-E
23. Odds ratio, risque relatif: estimation	A	I-E
24. Facteurs susceptibles d'influencer la puissance d'une étude	A	I-C
25. Calcul du nombre de sujets nécessaire	A	I-C
26. Connaissances de base sur les sources d'erreur et les biais	A	I-G
27. Connaissances plus avancées sur les biais.	B	I-G
28. Interaction / modification d'effet	B	I-G
29. Jugement de signification statistique, degré de signification et intervalle de confiance	A	I-C
30. Critère de jugement de causalité et niveau de preuve des différentes études.	A	III-B
31. Lister les nouvelles sources de données au service du progrès médical soutenu par la technologie (microbiome, exposome, phénomène et épigénome)	B	I-A
32. Modalités de la communication scientifique; processus de publication d'un article scientifique; revues prédatrices.	B	III-C
33. Evaluation de l'apport décisionnel des tests diagnostiques	A	I-F, II-A

■ **Item 323 – Analyser et utiliser les résultats des études cliniques dans la perspective du bon usage – analyse critique, recherche clinique et niveaux de preuve**

Intitulé connaissance	Rang	Chapitre
1. Connaître les notions d'EBM (définition, finalité) et d'évaluation thérapeutique (principes et champs d'application)	A	III-C
2. Argumenter l'évaluation d'un médicament ou d'une thérapeutique non médicamenteuse	A	II-C
3. Démarche et principe des méta-analyses d'essais	B	II-D
4. Définition, justifications indications et interprétation des études de pharmaco-épidémiologie	B	II-E
5. Connaître les notions d'effets placebo et nocebo dans les essais cliniques	A	II-C
6. Définition et rôle d'un médicament placebo	A	II-C
7. Principaux mécanismes et facteurs influençant l'effet placebo	B	II-C
8. Connaître les différentes sources d'informations thérapeutiques et savoir en apprécier la pertinence (littérature primaire, secondaire, tertiaire et quaternaire), recommandations de bonne pratique HAS, ANSM, INCA etc., fiche bon usage du médicament	A	III-C
9. Evoquer les liens d'intérêts et l'impact potentiel sur l'information médicale	A	III-C
10. Connaître les limites potentielles de l'information provenant de l'industrie pharmaceutique	A	III-C
11. Argumenter les niveaux de preuve des principales sources d'information	A	III-B
12. Taille de l'effet	A	I-E
13. Cadre de l'évaluation thérapeutique	A	II-C
14. Randomisation	A	II-C
15. Aveugle du patient et de la personne prenant en charge le patient et de l'évaluateur	A	II-C
16. Connaître les situations acceptables de simple insu	B	II-C
17. Assignment secrète dans les essais randomisés	A	II-C
18. Pertinence clinique	A	III-A
19. Transposabilité des résultats (validité externe ou extrapolation)	A	III-A
20. Validité interne, externe, cohérence externe	A	III-A
21. Connaître les principes de lecture critique d'un protocole d'essai clinique	A	III-C
22. Vérification de la comparabilité initiale des groupes dans un ECT	A	II-C
23. Connaître la notion de rapport bénéfice / risque	A	II-C
24. Typologie des essais randomisés	A	II-C
25. Analyse en intention de traiter et per protocole	A	II-C

PARTIE I
MÉTHODES

A DÉFINITION DES OBJECTIFS EN ÉPIDÉMIOLOGIE ET RECHERCHE CLINIQUE

J. Warszawski, en collaboration avec S. Novelli, A. Pastorello et A. Rouquette.

Connaissances concernées :

- 20-01-A : Définition et organisation de la recherche en santé
- 20-02-A : La méthodologie de la recherche en santé
- 20-04-B : Justification et problématique de recherche
- 20-05-B : Hypothèses de recherche
- 20-06-B : Méthodes adaptées à un objectif de recherche
- 20-16-A : Facteur de risque, facteur étiologique, facteur pronostique, déterminants de santé, facteur exposition
- 20-17-A : Population cible et source, échantillon
- 20-31-B : Lister les nouvelles sources de données au service du progrès médical soutenu par la technologie (microbiome, exposome, phénomène et épigénome)

Une étude en épidémiologie ou en recherche clinique a pour origine une ou plusieurs questions que se pose un chercheur, un clinicien, ou d'une manière plus générale, un ou plusieurs acteurs de santé publique, en l'absence de réponse satisfaisante compte tenu des informations disponibles dans la littérature scientifique ou autre. L'objectif de l'étude est donc de répondre à cette (ces) question(s). La mise en œuvre d'une étude (impliquant du temps, des moyens financiers et humains, et de l'énergie) ne sera justifiée qu'après **une revue approfondie de la littérature et une synthèse des données existantes, permettant notamment de formuler des hypothèses**, afin :

- De **justifier** et de **définir précisément le ou les objectifs** de cette nouvelle étude, **et**
- De **concevoir au mieux la méthodologie à mettre en œuvre**.

La définition du ou des objectifs est une **étape cruciale** qui va guider le chercheur tout au long de l'élaboration du protocole de son étude. **En Lecture Critique d'Article (LCA)**, **identifier précisément le ou les objectifs de l'étude** permet de juger de l'adéquation des méthodes utilisées pour l'atteindre, et donc d'identifier notamment les biais potentiels et le manque de puissance éventuel concernant les résultats présentés.

Ci-dessous sont décrites les 4 étapes essentielles à la définition opérationnelle d'un objectif, détaillées dans les chapitres suivants.

I - Préciser la question posée

Il s'agit de définir et formaliser précisément :

- **Le(s) critère(s) de jugement (« outcome »)**, noté(s) ici « **M** » (comme **m**aladie), mais qui correspond(ent), par extension, à un critère médical (maladie, symptôme, réponse à un traitement, décès, paramètre biologique...), comportemental, environnemental, en lien avec la santé.

- **Le(s) facteur(s) d'exposition éventuel(s), noté(s) ici « E » (= exposition)**
 - E est appelé **facteur de risque** si le niveau de E pour chaque sujet est recueilli mais n'est pas attribué de manière expérimentale dans le cadre de l'étude,
 - E est appelé **intervention expérimentale** si le niveau ou la catégorie de E est attribué pour chaque sujet à titre expérimental dans le cadre de l'étude, limité le plus souvent à des stratégies de prise en charge médicale ou préventive dont on cherche à montrer le bénéfice.
- **La population cible de l'étude, notée « P »** : celle à laquelle on souhaiterait **appliquer les résultats** (qui, quand, où dans l'idéal)
 - *NB* : La population cible est à distinguer de la **population source, notée « S »** comme sujet, qui est celle à laquelle on a accès pour réaliser en pratique l'étude, en définissant de manière précise les critères d'éligibilité, d'inclusion et de non-inclusion (en pratique : Qui/quels sujets, Quand, Où).

Remarques sur les notations : M et E peuvent être des critères quantitatifs (glycémie, âge, indice de masse corporelle...) ou qualitatifs binaires (statut vital : décédé/vivant ; malade/non malade ; sportif/non sportif) ou non binaires (âge : 15-29 ans/30-44 ans/45-54 ans, consommation de tabac : fumeur /ex-fumeur/ non-fumeur) – cf. Partie I.C.

L'acronyme **PICO ou PECO** est largement utilisé pour se rappeler de ces éléments essentiels à préciser lors de la formulation d'un objectif : **P**opulation (=P) / **I**ntervention expérimentale ou **E**xposition / **C**ontrôle, comparateur ou comparaison à... / **O**utcome, critère de jugement

Exemple : Evaluer l'efficacité du bisoprolol par rapport à un placebo sur la mortalité toute cause dans l'insuffisance cardiaque en Europe : M est la mortalité (à partir d'un critère binaire : décédé/vivant) / E est l'exposition qui est ici une intervention expérimentale (critère binaire : E+ = bisoprolol, E- = placebo), P est la population cible (patients atteints d'insuffisance cardiaque en Europe).

Avec la notation PECO :

- P = Patients atteints d'insuffisance cardiaque en Europe
- E = E+ = Bisoprolol
- C = E- = Placebo
- O = M+ = Mortalité toute cause



- Il existe **toujours au moins un critère de jugement M**
- Il y a **en général un seul critère de jugement principal** répondant à un **objectif principal** mais certaines études peuvent comporter plusieurs objectifs donc plusieurs critères de jugement (le plus souvent pour un objectif principal et des objectifs secondaires)
- Lorsque les objectifs sont exclusivement **descriptifs**, il n'y a **pas d'exposition**
- La population source doit dans l'idéal être **aussi proche que possible** de la population cible

II - Déterminer l'approche : observationnelle ou expérimentale

Dans une approche **observationnelle** : on ne modifie pas la nature et le niveau des facteurs E auxquels sont soumis les sujets inclus dans l'étude.

Exemple : comparer la tolérance rénale évaluée par la créatininémie (M = critère quantitatif) de 4 traitements anti-VIH (E à 4 catégories dont l'un des traitements est la catégorie de référence) prescrits selon les pratiques habituelles de soins chez les patients porteurs de l'infection à VIH (P).

Dans une approche **expérimentale** : les **sujets sont soumis** à une **intervention expérimentale E** qui n'aurait pas eu lieu, ou pas dans les mêmes conditions ou pas au même moment, s'ils n'étaient pas inclus dans l'étude (ce qui n'est pas toujours faisable ou éthique).

Exemple : évaluer l'efficacité d'un nouveau traitement anti-VIH (E+) administré dans le cadre d'une étude expérimentale, par rapport au traitement de référence (E-), en termes de diminution des complications rénales (M) chez les patients porteurs de l'infection à VIH (P).



Il ne faut pas confondre « **expérimentale** » et « **interventionnelle** »

- **Toutes les études expérimentales sont interventionnelles**, mais il existe des études interventionnelles qui ne sont pas expérimentales
- **La définition d'étude interventionnelle relève d'un cadre réglementaire plus large** (cf. Partie I.H : certaines études observationnelles peuvent être interventionnelles si on doit utiliser **spécifiquement pour l'étude** certaines méthodes de recueil de données que les participants n'auraient pas eu systématiquement en dehors de l'étude, ou modifier leur calendrier de suivi médical par exemple.... Dans ce cas l'étude pourra être qualifiée d'interventionnelle (**sans qu'elle soit expérimentale car il n'y a pas de modification de la nature ou des niveaux de E**)

III - Préciser l'unité de recueil (= d'analyse)

Une étude effectuée sur données **individuelles** consiste à collecter ou mesurer, **pour chaque sujet inclus dans l'étude**, le niveau ou statut de chaque critère de jugement **M** et d'exposition éventuelle **E**.

Exemple : pour étudier le lien entre risque d'accident cardio-vasculaire M et consommation d'alcool E, une étude sur données individuelles collecte pour chaque sujet sa consommation d'alcool d'une part, et la survenue ou non chez ce sujet d'accident cardio-vasculaire d'autre part, dont les modalités et la temporalité de la collecte dépendent du type d'étude (cf. infra).

Une étude sur données **groupées (= agrégées)** repose sur l'utilisation de **plusieurs sources de données car on ne connaît pas dans la même étude le statut individuel de M (et/ou de E)**. C'est le cas lorsqu'on estime par exemple le **taux d'incidence**¹ de M+ dans une population (par exemple, dans un département ou un pays...) où existe un **système d'enregistrement** qui recense tous les sujets développant l'événement **M+** au fur et à mesure du diagnostic (ex : registre de cancers, maladies à déclaration obligatoires). Ce système n'enregistre pas d'information sur les personnes qui ne développent pas l'événement (M-). Pour estimer le taux d'incidence de **M+** à partir des cas enregistrés pendant une période donnée dans cette population, on dispose des données du **recensement** qui fournit l'effectif de la population totale pendant la période considérée, ce qui permet d'obtenir le dénominateur (cf. Partie I.D). Ces taux d'incidence peuvent être comparés entre plusieurs populations ou dans le temps.

¹ Voir cours sur « mesures descriptives », partie I.D : le taux d'incidence est une mesure de fréquence de M+

Lorsqu'on dispose de différents systèmes d'enregistrement permettant d'estimer la fréquence de plusieurs indicateurs **M+** et **E+** dans des groupes de sujets, il est aussi possible de comparer les évolutions conjointes de ces fréquences dans le temps ou dans l'espace.

Exemple : estimer l'évolution du taux d'incidence des cancers standardisé sur l'âge et le sexe en France entre 1990 et 2020 à partir des registres de cancers, comparer ces taux d'incidence entre les différentes régions, étudier la corrélation entre l'évolution de ces taux d'incidence et l'évolution de la consommation de tabac et d'alcool sur la même période.

IV - Préciser la nature de l'objectif de l'étude

Il existe quatre grandes classes d'études selon que l'objectif est :

- à visée **descriptive**
- à visée **explicative en situation observationnelle**, souvent appelée étude **analytique**
- expérimental, c'est à dire visant à **démontrer l'effet d'une intervention expérimentale**, *i.e.* réalisée pour l'étude
- destiné à **évaluer les performances d'une méthode de classement** (test diagnostique) de M

IV.A - Etudes à visée descriptive

Intérêt : quantifier l'**ampleur** d'un phénomène, étudier ses **variations géographiques et temporelles**, étudier ses **variations en fonction d'autres indicateurs** caractérisant la même population, générer des hypothèses sur les facteurs susceptibles d'influencer la survenue de ces phénomènes, etc.

Approche : toujours **observationnelle**.

Unité de recueil :

- **Individus** : permettant l'**analyse de données individuelles**, consistant à obtenir le statut de M pour chaque individu d'un échantillon issu d'une enquête transversale au sein de la population ciblée pour l'estimation de prévalence, ou d'une enquête de cohorte à recrutement transversal initial pour estimer des taux d'incidence (*cf.* Partie I.B).
- **Groupes d'individus** : permettant l'**analyse de données groupées**, généralement au niveau d'une zone géographique administrative (pays, département, etc.) dont le nombre d'habitants est connu. Le recueil des données porte sur l'enregistrement de tous les cas M+ par des systèmes de surveillance épidémiologique (systèmes de collecte continue de survenue de certaines maladies) ou par le registre national des causes de décès, ou par des études plus spécifiques. Le statut M+/M- de chacun ou d'une partie de ses habitants n'est pas recueilli (*cf.* Partie I.B).

Indicateurs et principes d'analyse :

- Estimer des indicateurs de « fréquence » d'une pathologie ou d'un comportement de santé, comme un **taux d'incidence**, une **prévalence** ou une **distribution** d'une variable à plus de deux catégories, ou des indicateurs de type **moyennes ou médianes** pour des variables quantitatives (*cf.* Partie I.D).
- Comparer l'**évolution de ces indicateurs** au cours du temps ou dans l'espace à partir de données groupées sur des périodes ou des zones géographiques (période ou zone géographique = groupe d'individus).
- Étudier des **corrélations entre différents indicateurs** caractérisant des groupes de sujets issus de sources de données différentes (étude de corrélation écologique géographique ou temporelle, *cf.* Partie I.B).
- etc.

Principaux types d'étude :

- Etudes de surveillance épidémiologique
- Etudes de corrélations écologiques
- Enquêtes transversales et enquêtes de cohorte à visée représentative



- Les comparaisons des indicateurs de fréquence ou de niveau de M dans le temps et l'espace (entre pays par exemple) appartiennent au champ des **études descriptives** même si elles font l'objet d'analyses statistiques sophistiquées comportant des tests statistiques. Ces comparaisons ne permettent pas d'identifier des facteurs associés à ces différences temporelles ou spatiales. **Il n'y a pas d'exposition E.**
- Les études de **corrélation écologique** entre deux indicateurs caractérisant des pays par exemple, sont, malgré leur terminologie, habituellement considérées comme des études **descriptives** et non analytiques. En effet, ces corrélations portent sur des indicateurs qui caractérisent des populations et non des individus. Elles présentent ainsi d'importantes limites pour les interpréter. Par exemple, il peut arriver, en comparant plusieurs pays, que la proportion d'individus M+ soit d'autant plus élevée dans un pays que la proportion d'individus E+ l'est aussi, alors qu'à l'échelle d'un pays donné, la proportion des M+ soit la même chez les E+ et E-. L'association observée entre les indicateurs à l'échelle des pays (groupes) reflète le rôle caché d'un autre facteur impossible à prendre en compte sans disposer de données pour chaque habitant : c'est le « **biais écologique** ». Ce type d'étude reste néanmoins très utile pour générer ou renforcer des hypothèses justifiant de mener des études sur données individuelles (cf. Partie I.G).

IV.B - Etudes à visée explicative (études analytiques)

Intérêt : rechercher une association statistique entre un ou plusieurs facteurs E et M, dont on fait l'hypothèse qu'ils sont susceptibles de modifier le niveau de M ou le risque ou le délai de survenue de M+ si le critère est binaire.

L'exposition E est appelée **facteur de risque** d'une manière générique mais selon la nature de M, E peut être :

- Une **étiologie** d'une maladie (M)
- Un **facteur pronostique** de l'évolution défavorable (M) d'une pathologie
- Un **déterminant** de comportement (M) connu pour augmenter le risque d'une maladie
- Un **facteur socio-économique** structurel susceptible de favoriser des pratiques de soins non optimales (M)
- etc.

Approche : toujours **observationnelle**.

Unité de recueil : données **individuelles** obtenues dans le cadre d'enquêtes de **cohorte**, d'enquêtes **cas-témoins**, ou d'enquêtes **transversales** (cf. Partie I.B).

Indicateurs et principes d'analyse : Ces études **reposent obligatoirement sur des comparaisons de groupes de sujets** (M+/M- ou E+/E- selon le schéma d'étude), et consistent à :

- Tester **l'association statistique** entre un ou plusieurs facteurs E (non attribués à titre expérimental) et le risque de survenue, le délai de survenue ou le changement de niveau d'un critère de jugement M (cf. Partie I.C)
- Estimer des **mesures d'association**, comme un odds-ratio, un risque relatif, un hazard ratio, une différence de risque, une différence de moyenne (cf. Partie I.E)

Principaux types d'étude :

- Etudes étiologiques
- Etudes pronostiques (cf. Partie II.B)



- Ces études **ne permettent pas d'affirmer la nature causale** des associations statistiques observées. **Un facteur de risque E est un facteur associé statistiquement à M.** C'est pourquoi on peut parler de facteur de risque même si on le présente sous forme de **facteur protecteur**, car le sens dépend de la catégorie de référence choisie : il suffit d'inverser le codage de E pour inverser la mesure de l'association (ou en changer le signe). Dans tous les cas, déterminer un facteur de risque (ou son miroir appelé facteur protecteur) revient à mettre en évidence une association statistique dont on donne le sens, mais qui ne préjuge pas de la nature causale ou non de cette relation.
- Les expositions E étudiées sont des facteurs pour lesquels ont fait l'hypothèse d'une influence directe sur la survenue de M. Ces facteurs d'exposition E sont à distinguer des « **déterminants de santé** » qui peuvent influencer l'état de santé des populations, sans nécessairement être des causes directes de comportement ou de maladie, et qu'on peut chercher à prendre en compte dans certaines analyses.

IV.C - Etudes expérimentales

Intérêt : démontrer l'**effet** (positif ou négatif) d'une intervention expérimentale **E** réalisée pour l'étude sur un critère de jugement **M** (état de santé, comportement, qualité de vie...)

Approche : **expérimentale**

Unité de recueil :

- Données **individuelles** si l'intervention expérimentale est menée à l'échelle individuelle
- Données **groupées** si l'intervention expérimentale est menée à l'échelle populationnelle (plus rare)

Indicateurs et principes d'analyse : ces études **reposent obligatoirement sur la comparaison des sujets** :

- ✓ **Soumis (E+) et non soumis (E-)** à une intervention expérimentale, ou soumis à deux modalités différentes d'une intervention expérimentale.
- ✓ **Appartenant à des populations ayant reçu (E+) ou non (E-)** une intervention expérimentale, ou à 2 deux modalités différentes d'une intervention expérimentale.
- ✓ **Avant et après** une intervention expérimentale E+ (sans groupe contrôle).

Dans les deux premiers cas, si les modalités E+/E- sont attribuées par tirage au sort, il s'agit d'un essai **randomisé** individuel, ou en cluster. L'analyse consistera à :

- Tester l'**association statistique** entre l'intervention expérimentale (**E**) et un événement (**M**) qu'elle vise à modifier.
- Estimer des **mesures d'association** entre l'intervention expérimentale (**E**) et un événement (**M**) qu'elle vise à modifier.

Principaux types d'étude :

- **Essais randomisés thérapeutiques** (cf. Partie II.E), de **stratégies diagnostiques, préventives, etc.**
- **Essais randomisés de prévention en population.**
- **Etudes quasi-expérimentales** dans laquelle l'intervention expérimentale n'est pas attribuée par tirage au sort, avec ou sans groupe contrôle, mais où l'indicateur d'efficacité potentielle est mesuré avant et après (effets **avant-après**) une intervention expérimentale en population

(**enquêtes transversales répétées**) ou dans un échantillon (**enquête de cohorte**), ou bien dans différentes zones géographiques (effets **ici-ailleurs**)



- *Les essais randomisés contrôlés, correctement menés, sont les seules études permettant de conclure à la nature causale d'une association mise en évidence entre l'intervention expérimentale E et le critère de jugement principal M.*
- *Le schéma de l'essai randomisé s'applique classiquement à l'évaluation de l'effet d'une stratégie thérapeutique, préventive ou diagnostique administrée au niveau **individuel** (en recherche clinique). Il peut s'appliquer à l'évaluation d'une intervention de santé publique à l'échelle **populationnelle**, comme une campagne de promotion active de vaccination par exemple : on attribue par tirage au sort les régions qui recevront ou non cette campagne de promotion et on compare ensuite la proportion de sujets vaccinés (M+) entre les régions ayant reçu ou non l'intervention expérimentale.*
- *On peut évaluer l'effet d'une stratégie thérapeutique E sur le pronostic d'une maladie M dans le cadre d'une enquête observationnelle sur données individuelles : il s'agit alors d'une étude analytique à visée explicative qui recueille l'exposition ou non à un traitement E+ dans le cadre du suivi médical de routine : ce traitement n'a pas été attribué pour l'étude (cf. Partie II.E par exemple). Les conclusions ne **permettront pas de démontrer** l'effet causal éventuel.*

IV.D - Etudes visant à évaluer les performances d'une méthode de diagnostic ou de dépistage (méthode de classement de M)

Intérêt : Evaluation des **performances** d'un ou plusieurs tests **T** pour le diagnostic ou le dépistage de **M** permettant classer correctement en M+ et M- (classés par une **méthode de référence**) ou **comparaison des performances de plusieurs méthodes de classement** entre elles.

Ce type d'étude s'applique plus généralement à l'évaluation des performances d'une méthode de classement d'un événement ou d'une caractéristique (par exemple les performances d'un questionnaire standardisé pour classer le plus correctement possible la consommation de tabac actuelle ou passée des participants à une étude). D'une manière générale, une méthode est d'autant plus « performante » qu'elle minimise le nombre de faux positifs et de faux négatifs (cf. Partie I.F).

Approche : **observationnelle** ; le plus souvent interventionnelle (mais **jamais expérimentale** car il n'y a pas de E).

Unité de recueil : données **individuelles** dans le cadre d'une enquête **transversale** (avec éventuellement suivi longitudinal dans certains cas).

Indicateurs et principes d'analyse :

- Evaluation des **performances diagnostiques** d'une méthode à évaluer (**T**) par rapport à une méthode référence (référence standard, **M**) : sensibilité, spécificité, valeurs prédictives positives et négatives, rapports de vraisemblance, aire sous la courbe ROC
- Comparaison de plusieurs méthodes (**T**) entre elles

Principaux types d'étude :

→ **Etudes diagnostiques** (cf. Partie II.C)



- Dans ce type d'étude, il n'y a **pas d'exposition** (la ou les méthodes T à évaluer ne sont pas des expositions). Ce type d'étude n'est donc **jamais expérimental**, et relève bien d'une approche observationnelle.
- En revanche, ce type d'étude est **le plus souvent interventionnelle (sur le plan réglementaire)** car chaque participant est soumis pour l'étude à la méthode de référence ET à la méthode à évaluer, ce qui est rarement le cas dans la pratique de routine.
- Ces études, dites diagnostiques, ne doivent pas être confondues avec des études (expérimentales ou analytiques) visant à évaluer l'efficacité d'une nouvelle démarche diagnostique (= E) d'une maladie sur son pronostic (=M)

Exemple : la réalisation d'un scanner thoracique systématique = E, chez des personnes de 50 à 74 ans consommant du tabac depuis plus de 10 ans, permet-elle une réduction de la mortalité par cancer du poumon = M ? Ici on ne cherche pas évaluer les performances diagnostiques du scanner thoracique mais si un diagnostic systématique par cette méthode serait efficace pour réduire la mortalité

- L'enquête **transversale** est le schéma classique des études diagnostiques. Mais une période de suivi longitudinal peut être nécessaire pour collecter le classement correct de M en différé (par exemple attendre le décès pour l'autopsie).

V – Rédiger un protocole

En France, la recherche clinique est réglementée par des textes de loi (cf. Partie I.H).

Pour toute étude, un protocole doit être rédigé avant le début de l'étude.

Ce protocole ne doit pas être modifié en cours d'étude. Mais il peut arriver qu'une modification du protocole soit nécessaire en cours d'étude (extension de la période de recrutement ou au contraire arrêt prématuré, demande d'une analyse intermédiaire non prévue, etc...). Ces modifications doivent être justifiées scientifiquement, et faire éventuellement l'objet d'un **amendement réglementaire** selon le type de recherche concernée (cf. Partie I.H). Les modifications éventuelles doivent être indiquées dans les rapports et articles présentant les résultats de ces travaux.

Le protocole doit préciser les différents points qui sont décrits dans ce chapitre :

1. Le contexte, les résultats déjà connus de la littérature scientifique, etc.
2. Les hypothèses et questions visées par l'étude
3. Le ou les objectifs
4. La méthodologie de la recherche (cf. Fiches 1 et 2) :
 - a. Type d'étude (essai randomisé, étude de cohorte, étude de corrélation écologique, etc.)
 - b. Critère de jugement principal +/- secondaires
 - c. Eventuellement type d'expositions étudiées ou intervention expérimentale éventuelle
 - d. Critères d'inclusion et de non inclusion
 - e. Modalités pratiques de sélection et de recueil des données
 - f. Calcul du nombre de sujets nécessaire
 - g. Plan d'analyses statistiques prévues

Tableau 1 : Exemples de sources de données disponibles pour les études biomédicales

Type de données	Natures des données
Données disponibles dans le SNDS	Regroupe les données <ul style="list-style-type: none"> · de l'assurance maladie (base SNIIRAM), · des hôpitaux (PMSI), · des causes médicales de décès (base CépiDC), · des données relatives au handicap (MDPH) La mise à disposition de ces données vise à favoriser les recherches ou évaluations ayant un caractère d'intérêt public
Données de soins disponibles dans entrepôts de données de santé	Examen clinique, prescriptions, données médico-administratives du PMSI, imagerie, biologie, anatomopathologie, données des dispositifs médicaux (ECG, etc.), comptes-rendus, etc.
Données issues de biologie moléculaire	Microbiome (ensemble des micro-organismes qui colonisent le corps), métabolome (ensemble des métabolites retrouvés dans un échantillon biologique), épigénome (ensemble des modifications chimiques sur l'ADN et ses protéines associées), etc.
Données d'objets connectés	Montres, balances, smartphones, etc.
Génétique	Génome : ensemble de l'information génétique Phénomène : ensemble des caractères observables
Données environnementales	Météorologiques, pollution atmosphérique, radioactivité, etc. Exposome : ensemble des expositions environnementales au cours de la vie (environnement physique et mode de vie)

VI – Organisation de la recherche en santé en France

Les recherches en santé concernent les études scientifiques visant à développer de nouvelles connaissances en biologie et médecine et dans d'autres domaines ayant une application en santé (sciences du médicament, sciences humaines et sociales, nouvelles technologies, etc.), et à évaluer l'efficacité, l'utilité et la mise en œuvre de stratégies ou d'interventions susceptibles d'améliorer la santé, dans toutes ses dimensions.

En France, les grandes orientations stratégiques, l'organisation et les moyens sont déterminés par les **ministères de la santé et de la recherche**. La répartition de ces moyens est programmée par des agences généralistes non ciblées exclusivement sur la santé² ou plus spécifiques³. La recherche est organisée par des laboratoires de recherche gérés par des opérateurs, dont les Centres Hospitalo-Universitaires (CHU) et l'Institut National de la Santé et de la Recherche Médicale (INSERM) exclusivement dédiés à la recherche médicale, ou d'autres organismes non spécifiques de la recherche médicale⁴, ainsi que par des fondations reconnues d'utilité publique⁵. L'agence nationale « Santé Publique France », établissement public administratif, acteur principal de la surveillance épidémiologique en France, apporte aussi une contribution aux recherches en santé.

² ANR = Agence Nationale de la Recherche

³ Notamment : l'IRESP (Institut pour la Recherche en Santé Publique), ANRS-MIE (Agence Nationale de Recherche sur le sida et les hépatites virales | MIE) ; InCA (Institut National du Cancer)

⁴ Notamment : CNRS = Centre National de la Recherche Scientifique ; INRAE = Institut National de la Recherche Agronomique et de l'Environnement) ; Universités...

⁵ Notamment : Institut Gustave Roussy, Institut Pasteur de Paris, Institut Curie, Institut Pasteur de Lille ; CEPH (Centre d'étude du polymorphisme humain)



DÉFINITION DU SCHEMA D'ÉTUDE SELON LES OBJECTIFS

P(I)ECO-T = **P**opulation / **E**xposition ou **I**ntervention expérimentale / **C**ontrôle, comparateur ou comparaison à... / **O**utcome, critère de jugement / **T**ype d'étude, design

POPULATION CIBLE

Critères d'éligibilité **S** (= critères d'inclusion et de non-inclusion) que tous les sujets de l'étude doivent présenter

- **Période d'éligibilité** = période pendant laquelle les critères d'inclusion/non-inclusion¹ sont réunis

VARIABLES D'INTERET

- **Critère de jugement M** : il peut être unique (principal) ou non (en général un critère principal **M** +/- des critères secondaires **M'**) ; il peut être qualitatif (binaire, nominal, ordinal), quantitatif, censuré ou non ; il peut être simple ou composite (défini à partir d'un ensemble de critères)
- **Exposition / intervention expérimentale E** = facteur susceptible de modifier le risque ou le délai de survenue de M+ ou le niveau de M : elle peut être unique ou non / principale ou non / de nature qualitative ou quantitative / stable ou non au cours du temps

⚠ *Si E = intervention expérimentale : préciser la nature et les modalités d'attribution de E aux sujets de l'étude (randomisation ou choix raisonné / au niveau individuel ou populationnel)*

- **Test diagnostique T** : un test T à évaluer à l'aide d'une méthode de référence capable de classer correctement chaque sujet M+/M-, ou comparaison de plusieurs tests entre eux (T1, T2, etc.)

SCHEMA DE L'ÉTUDE²

Déterminé par les modalités : (1) de sélection/inclusion des unités analysées, (2) de recueil des informations pour chacune de ces unités concernant au minimum **M**, +/- **E**. Ces modalités dépendent :

- **De l'objectif (but de l'étude)** : descriptif / explicatif (= analytique) / d'évaluation d'une intervention expérimentale / d'évaluation des performances d'une méthode diagnostique
- **De l'approche méthodologique** : observationnelle / expérimentale
- **Du type d'unités de recueil/d'analyse** : individu / groupe (données agrégées)
 - Observationnelle sur données groupées : **analyse des données de surveillance épidémiologique, étude de corrélation écologique**
 - Observationnelle sur données individuelles : **enquête de type transversale, cohorte, ou cas-témoins**
 - Expérimentale (ou quasi-expérimentale³) sur données individuelles ou sur données groupées selon que l'intervention E évaluée est menée à l'échelle des individus ou d'une population : **le prototype de l'étude expérimentale en recherche biomédicale est l'essai randomisé**

⚠ NOTES

¹ les **critères de non-inclusion** sont les critères faisant que les patients ne peuvent pas être inclus dans l'étude, les **critères d'exclusion** définissent, parmi les sujets inclus dans l'étude, lesquels devront être exclus de l'analyse. Cette distinction est absente en langue anglaise où tous ces critères sont dénommés « *exclusion criteria* ».

² On trouve souvent dans les protocoles une section décrivant le « **schéma expérimental** », parfois appelé « **design** » ou « **plan expérimental** » (à ne pas confondre avec « étude expérimentale »), dans lequel on décrit le type d'étude et éventuellement les détails de sa réalisation en pratique. Cette expression est utilisée qu'il s'agisse d'études observationnelles ou expérimentales).

³ On parle d'études **quasi-expérimentales** lorsque E est attribuée pour une étude en dehors d'un essai randomisé : suivi des sujets ayant reçu une intervention expérimentale sans groupe contrôle, ou sans attribution randomisée (essai de phase II, cohorte interventionnelle), comparaison de prévalence de M+ avant/après une intervention E (enquêtes transversales répétées), ou entre deux zones géographiques (enquêtes transversales ou de cohortes ici/ailleurs).

Type d'étude	Diagnostique	Descriptive	Explicative = analytique	Expérimentale ou quasi-expérimentale
Approche méthodologique = modalité d'attribution de E	Evaluation des performances d'une méthode T de classement de M	Estimation de « fréquence de M » ou étude des Variations conjointes de plusieurs indicateurs dans temps ou espace	Recherche de facteur(s) E de risque de M (étiologiques, pronostiques, comportementaux...)	Evaluation de l'effet d'une intervention E, réalisée pour l'étude, sur le risque de M («quasi » si E non randomisée)
Approche réglementaire	OBSERVATIONNELLE			
Unité de recueil/analyse de M	Individus	Groupe	Individus	Individus ou Groupes
Unité de recueil/analyse de E	Test ≠ E	Pas de E ou « E groupé »		E= intervention
Système de surveillance épidémiologique	Pas de E	E groupé	Individus	Groupes
Etude de corrélation écologique	Enregistrement de cas M+	M et E venant de sources ≠ 2		Individus
Enquête transversale	≥ 2 méthodes différentes de classement de M	Unique ou répétée	Unique ou répétée	Répétée avant/après une intervention E attribuée pour l'étude
Enquête de cohorte	+/- si besoin d'un suivi longitudinal pour classer a posteriori M	Avec recrutement transversal	recrutement transversal ou exposé/non exposé E+/E-	E+ et E- attribué de manière raisonnée
Enquête cas-témoins			Eventuellement nichée dans un autre étude (cohorte par ex)	
Essai randomisé				E+ et E- attribué par tirage au sort (essai en cluster)

¹ = facteur susceptible de modifier le risque ou le délai de survenue de M : l'étude vise à mettre en évidence une association statistique avec M sans pouvoir en affirmer la nature causale

² = classification en « étude descriptive » car permet de générer des hypothèses sur des facteurs potentiellement associés à un état de santé. Mais ces études présentent des limites importantes pour interpréter une relation entre E et M mise en évidence au niveau de groupes, car ne permettent pas de s'assurer que ce sont les sujets exposés qui sont les plus à risque de développer M que les sujets non exposés



MODALITÉS D'ÉCHANTILLONNAGE ET RECUEIL DE DONNÉES

Date de conception de l'étude : date où les objectifs de l'étude ont été définis et le protocole finalisé.

À DÉCRIRE	SIGNIFICATION
Population cible	Qui, où, quand dans l'idéal Population à laquelle on souhaiterait appliquer les résultats de l'étude
Population source Critères d'éligibilité Critères d'inclusion/non-inclusion	Qui, où, quand en pratique Liste, fichier, structure ou lieu (sites/centres d'inclusion) permettant d'identifier et sélectionner des sujets éligibles pour le recueil de données sur M (critère de jugement) +/- E (exposition), au cours d'une période donnée (période/dates d'inclusion), à partir des critères d'éligibilité : <ul style="list-style-type: none"> • Critères d'inclusion : caractéristiques que doivent présenter les sujets pour être inclus (=critères positifs) • Critères de non-inclusion : caractéristiques que ne doivent pas présenter les sujets éligibles a priori pour être inclus (= critères négatifs) ≠ critères d'exclusion Parmi ces critères doivent figurer le consentement éclairé ou la non-opposition pour la participation selon le type réglementaire de l'étude. <i>NB (1) : critères d'exclusion</i> : caractéristiques des sujets initialement inclus dans l'étude qui devront être exclus de l'analyse post-inclusion. <i>NB (2) En langue anglaise</i> : les critères de non inclusion et les critères d'exclusion sont dénommés « <i>exclusion criteria</i> » sans distinction
Modalités de recrutement des sujets éligibles	Sélection : <ul style="list-style-type: none"> • Tirage au sort ou recrutement exhaustif dans une liste/fichier disponible • Sélection consécutive au fur et mesure de leur présentation dans les centres et pendant la période d'inclusion ; exhaustif ou par tirage systématique (1 sur 2 ou sur 3, etc...) • Sélection de convenance / choix raisonné (selon disponibilité et accessibilité) • Appel à volontaires <p><i>⚠ Il est rare que tous les sujets soient recrutés exactement au même instant, le recrutement est en général échelonné sur une période de temps (période d'inclusion), même dans le cas d'une sélection à partir d'un fichier initial, car il faut ensuite contacter les sujets sélectionnés, obtenir leur consentement et/ou recueillir les données</i></p>
Effectif à inclure	Calcul à effectuer avant la réalisation de l'étude en fonction de critères de précision/puissance souhaitée (= calcul du nombre de sujets nécessaires , de la taille d'échantillon nécessaire)
Participants/sujets inclus	Sujets éligibles effectivement inclus c'est à dire ayant pu être contactés ou dont le dossier a été retrouvé selon les modalités de l'étude, ET répondant à l'ensemble des critères d'inclusion et de non-inclusion
Recueil de données sur M et E dans les études observationnelles	Données spécifiques à l'étude (uniquement en cas de recrutement prospectif) : nécessite obligatoirement un contact direct avec le patient ou l'entourage (ex. questionnaires administrés au sujets et/ou entourage / examens cliniques / examens complémentaires, etc.) Données non spécifiques à l'étude recueillies de manière si possible standardisée : <ul style="list-style-type: none"> • Dans les dossiers médicaux, administratifs, etc. • Par fusion de bases de données (e.g., causes médicales de décès, registres) sur identifiant commun
Attribution de E dans les études expérimentales	Dans un essai randomisé, le statut de l'intervention expérimentale E est attribué par tirage au sort aux sujets éligibles, constituant ainsi les bras de l'essai (groupes comparés).

Suivi longitudinal des sujets inclus	<p>Permet le recueil de données correspondant à une période postérieure à l'inclusion pour chaque sujet inclus, jusqu'à une date de point ou la date de fin de l'étude, avec au minimum :</p> <ul style="list-style-type: none"> - Recueil de l'évolution du statut de M avec la date de survenue du changement de statut, en particulier la survenue de l'événement M+ s'il s'agit d'un critère binaire - Recueil de la date des dernières nouvelles en l'absence de survenue de M+ : date de décès, date de perte de vue, date de point ou date de fin de l'étude <p><i>NB : la date de point est une date butoir pour l'analyse statistique des données (spécifique des études avec suivi longitudinal) à laquelle tous les efforts seront faits pour connaître l'état de chaque sujet inclus vis-à-vis de M et au-delà de laquelle les informations recueillies ne seront pas prises en compte dans les analyses.</i></p>
Temporalité	<p>La période de recrutement des sujets peut être postérieure ou antérieure à la date de conception de l'étude :</p> <ul style="list-style-type: none"> • Postérieure = Recrutement prospectif : permet un contact direct si besoin avec les sujets inclus pour planifier le recueil des données à l'inclusion, et éventuellement postérieur à l'inclusion dans les enquêtes de cohorte • Antérieure = Recrutement rétrospectif : ne permet plus de contacter directement certains sujets inclus (si décédés ou perdus de vue au moment de l'étude), ni de planifier le recueil des données en fonction des objectifs actuels <p><i>△ Il ne faut pas confondre la durée du suivi dans une enquête de cohorte ou un essai clinique par exemple (de la date d'inclusion à la date de point ou de fin d'étude, éventuellement variable d'un sujet à l'autre) et la durée du recrutement qui peut s'étendre sur une période plus ou moins longue et concerne toutes les études (y compris cas-témoins ou transversales, i.e. sans suivi).</i></p> <p>Le recueil de données peut être prospectif ou rétrospectif :</p> <p>Recueil prospectif : données recueillies à l'inclusion, et après l'inclusion en cas de suivi longitudinal, dans les situations suivantes :</p> <ol style="list-style-type: none"> (1) Au fur et à mesure d'un recrutement prospectif, ce qui permet de planifier les données à recueillir et les conditions de recueil. (2) Dans le cadre d'une (sous-)étude nichée au sein d'une autre étude (de type cohorte prospective), d'un registre, ou dans un système collectant en routine des données sur les sujets concernés par l'étude (type base de patients hospitaliers). L'étude initiale ou le dispositif de routine constituent une partie des sources de données de la nouvelle étude, si besoin complétées d'un recueil prospectif ou rétrospectif de données complémentaires. <p><i>△ Dans les études transversales, il est inhabituel de parler de recueil prospectif des données présentes à l'inclusion, même lorsque le recrutement des sujets est prospectif. On parle quelques fois de recueil contemporain de l'inclusion.</i></p> <p>Recueil rétrospectif : concerne tous les autres cas où il est nécessaire de faire appel à la mémoire du sujet/son entourage ou de retourner collecter dans des documents (dossiers médicaux, administratifs, fichiers etc.) après la sélection des sujets, comme c'est presque toujours le cas pour les études cas-témoins. Le plus souvent, la disponibilité de ces données n'est pas homogène ou systématisée, mais on en prévoit des modalités de recueil <i>a posteriori</i> pour l'étude le plus standardisées possible.</p> <p><i>△ On ne définit pas la nature rétrospective ou prospective du recrutement ou du recueil des données à partir de la date à laquelle on effectue l'analyse statistique. Une analyse statistique est par définition toujours postérieure à l'obtention des données d'inclusion dans une enquête cas-témoins ou transversale, et ou de suivi dans une enquête de cohorte.</i></p>
Echantillon final d'analyse	<p>Echantillon des sujets inclus dont les données sont disponibles pour l'analyse, après l'application d'éventuels critères d'exclusion, par exemple pour :</p> <ul style="list-style-type: none"> • Retrait du consentement avec opposition à l'utilisation des données déjà collectées • Données manquantes sur des critères essentiels (en l'absence de procédure d'imputation prévue ou non pour remplacer les données manquantes).

VII – Entraînement : questions individuelles

Q1 – QRU – Cochez la proposition correcte parmi les suivantes :

- A. Une étude observationnelle n'est jamais expérimentale
- B. Une étude interventionnelle est toujours expérimentale
- C. Une étude observationnelle ne peut jamais être qualifiée d'étude interventionnelle
- D. Dans une étude expérimentale, l'intervention est le critère de jugement
- E. Le test à évaluer dans une étude diagnostique est le critère de jugement

Q2 – QRM – Cochez la ou les propositions correctes parmi les suivantes :

- A. Une étude analytique est toujours observationnelle
- B. Une étude explicative est toujours expérimentale
- C. Une étude observationnelle est toujours descriptive
- D. Une étude sur données groupées ne peut jamais être analytique
- E. Une étude sur données individuelles peut être expérimentale

Énoncé (*inventé, mais méthode inspirée de l'étude EnClass 2022 de Santé Publique France*) : Des auteurs suggèrent que l'échec scolaire pourrait retentir sur la santé mentale des adolescents. Pour étudier la relation entre l'auto-perception de la réussite scolaire et la santé mentale chez les lycéens de Paris en 2024, un échantillon de 10 lycées publics a été tiré au sort en mars 2024 parmi les 115 lycées publics de Paris. Le protocole prévoyait de convoquer tous les élèves, dans chaque lycée sélectionné, pour répondre à un auto-questionnaire en ligne administré sur place au lycée dans une salle surveillée, en l'absence d'opposition des parents des jeunes mineurs, ou de refus de l'élève lui-même, informés par courrier. Un consentement était demandé au début du questionnaire en ligne à l'élève lui-même.

Le questionnaire comportait une échelle validée pour évaluer la santé mentale, des questions sur la vie personnelle et familiale, les activités sportives et de loisirs, les consommations de substances psychoactives actuelles et passées, et l'auto-évaluation du niveau scolaire actuel, allant de 1 (très faible) à 5 (très élevé), à partir des résultats des deux premiers trimestres de l'année en cours.

Pour réaliser une étude complémentaire, le niveau scolaire a été également évalué de 1 à 5 pour chaque élève à partir des résultats des deux premiers trimestres de l'année fournis par le proviseur. Tout questionnaire sans réponse à plus d'un tiers des questions était considéré comme inexploitable

Les élèves convoqués ont rempli le questionnaire entre avril et juin 2024. L'analyse des données a été réalisée en septembre 2024. Une opposition active des parents a concerné 9% des élèves mineurs et un refus de participation a été indiqué par 3% des élèves. Au total, 6% des questionnaires étaient inexploitable selon la définition ci-dessus.

Q3 – QRM – Cochez la ou les propositions justes parmi les suivantes :

- A. La population cible est l'ensemble des lycéens de Paris
- B. La population source est l'ensemble des élèves des 10 lycées sélectionnés
- C. La non-opposition des parents est un critère d'inclusion
- D. Le refus d'un élève constitue un critère d'exclusion
- E. Aucun critère d'exclusion n'est indiqué dans cet énoncé

Q4 – QRM – Cochez la ou les propositions vraies parmi les suivantes :

- A. Le recrutement des élèves est prospectif
- B. Le recueil des données est rétrospectif car l'analyse a été réalisée 3 mois après ce recueil
- C. L'étude comporte un suivi longitudinal entre avril et juin 2024
- D. Les données recueillies reposent exclusivement sur les auto-questionnaires
- E. Il s'agit d'une étude sur données individuelles

Q5 – QRU – Cochez la proposition vraie parmi les suivantes :

- A. Il existe un seul critère de jugement principal
- B. Il existe plusieurs expositions principales
- C. Il existe au moins une exposition principale binaire
- D. L'exposition principale est le fait d'être un lycéen
- E. L'auto-évaluation de la réussite scolaire est le critère de jugement principal

Q6 – QRM – Cochez la ou les propositions vraies parmi les suivantes :

- A. L'objectif principal est analytique
- B. L'objectif principal est descriptif
- C. Il s'agit d'une étude expérimentale
- D. Il s'agit d'une étude observationnelle
- E. Il s'agit d'une étude interventionnelle

CORRECTION

Q1 - Réponse : A. A vraie car une étude observationnelle s'oppose à une étude expérimentale : dans le premier cas, les expositions *E* observées ne sont pas introduites ou modifiées pour l'étude ; dans le deuxième cas, une partie ou tous les participants de l'échantillon sont soumis à une exposition qu'ils n'auraient pas eu, ou pas sous cette forme s'ils n'avaient pas participé à l'étude, appelée alors intervention expérimentale ; **B et C fausses** car une étude observationnelle n'est jamais expérimentale mais peut être qualifiée d'étude interventionnelle sur le plan réglementaire si, par exemple, l'exposition étudiée est : a) non attribuée ou modifiée pour l'étude, b) recueillie de manière systématique alors que ce n'est pas systématique dans la prise en charge courante des participants et c) ce recueil implique des risques ou contraintes même minimes pour les sujets, exemple : radiographie, scanner ou IRM sans injection ; **D fausse** car l'intervention évaluée a le statut d'une exposition, mais attribuée pour l'étude ; **E fausse** car dans l'évaluation des performances diagnostiques, on n'étudie pas d'association entre *E* et *M* mais les performances de classement d'un test *T* à l'aide d'une méthode de référence qui permet de déterminer le statut vis-à-vis de *M*.

Q2 - Réponses : A, D et E. A vraie car aucune exposition étudiée n'est attribuée ou modifiée pour l'étude ; **D vraie** car ce type d'étude comporte potentiellement des « biais écologiques » ne permettant pas d'interpréter la relation entre *E* et *M* au niveau individuel ; **E vraie** car les études expérimentales peuvent être sur données individuelles, comme la majorité des essais randomisés, ou sur données groupées comme les essais randomisés en cluster par exemple ; **B fausse** car une étude explicative est synonyme d'étude analytique ; **C fausse** car une étude observationnelle peut être descriptive, analytique ou viser à étudier les performances diagnostiques d'un test *T*.

Q3 – Réponses : A et C. A vraie car la population cible est bien l'ensemble des lycéens de Paris comme indiqué dans la première phrase de l'énoncé ; **C vraie** car pour passer le questionnaire, il ne faut pas avoir reçu une opposition parentale ; **B fausse** car les 10 lycées inclus ont été tirés au sort parmi l'ensemble

des lycées publics de Paris éligibles ; la population source est donc l'ensemble des élèves scolarisés dans un des 115 lycées publics à Paris parmi lesquels ont été tirés au sort les 10 lycées ; **D fausse** car c'est un critère de non inclusion ; **E fausse** car 3% des questionnaires étaient inexploitables, ce qui constitue un critère d'exclusion.

Q4 – Réponses : A et E. A vraie car le recrutement a lieu après l'élaboration du protocole qui a permis de planifier l'ensemble des modalités de recueil actif des données ; **E vraie** car on collecte des données sur le statut de E= auto-perception de la réussite scolaire, et M =santé mentale pour chaque sujet) ; **B fausse** car une analyse statistique est toujours postérieure au recueil des données, ce qui ne signifie pas que l'étude est rétrospective : ici le critère de jugement est évalué par une échelle de santé mentale contemporaine de l'inclusion dans l'étude, et l'auto-évaluation des résultats porte sur la période actuelle ou récente ; **C fausse** car il n'y a pas de recueil de données pour ces sujets sur la période postérieure à l'inclusion (en particulier pas de données sur l'évolution de la santé mentale et/ou de la réussite scolaire), ce qui définirait un suivi longitudinal, spécifique à une enquête de cohorte ; c'est une étude transversale, dont la période de recrutement s'étale d'avril à juin 2024 ; **D fausse** car les résultats scolaires obtenus dans le cadre de la scolarité et non recueillis spécifiquement pour cette étude sont également fournis par le proviseur pour être analysés dans l'étude.

Q5 – Réponse : A. A vraie (la santé mentale) ; **B fausse** car l'exposition principale est l'auto-perception de l'échec scolaire (voir objectif), l'échec scolaire évaluée à partir des résultats fournis par le proviseur est une exposition secondaire, pour réaliser une étude complémentaire, par exemple pour étudier, à partir de la corrélation entre auto-perception de l'échec et résultats obtenus, si les élèves ayant une mauvaise santé mentale ont une exigence plus élevée vis-à-vis de leurs propres résultats scolaires ; **C fausse** car E= auto-perception de l'échec scolaire, est recueilli selon une échelle de 1 à 5 ; **D fausse** car c'est un critère d'éligibilité ; **E fausse** car c'est l'exposition principale.

Q6 – Réponses : A et D. A et D vraies car c'est une étude visant à étudier une association entre M et E sur données individuelles dans une enquête observationnelle ; **B fausse** car vise à étudier une association entre E et M dans une étude observationnelle ; **C fausse** car on ne modifie pas pour l'étude la réussite scolaire ; **E fausse** car aucun examen spécifique n'est nécessaire pour mesurer E ou M, obtenus par un auto-questionnaire (cf. Partie I.H)

B SCHEMA (DESIGN) DES ETUDES

J. Warszawski, avec la collaboration de S. Novelli, A. Pastorello, et A. Rouquette

Connaissances concernées :

- 20-07-B : Modalités d'échantillonnage
- 20-08-B : Recueil des données
- 20-09-B : Pertinence des données recueillies
- 20-18-A : Principaux types d'études observationnelles
- 20-19-B : Etudes au schéma moins classique

Une fois le ou les objectifs définis clairement, le choix du schéma (design) de l'étude à mettre en œuvre pour la collecte de données est basé sur des **critères éthiques** (est-ce éthique d'exposer expérimentalement des sujets à l'exposition étudiée ?) **et de faisabilité** (doit-on concevoir et réaliser une enquête *ad hoc* ou analyser des données déjà disponibles ? Le critère de jugement M est-il rare ou fréquent dans la population étudiée ? Peut-on trouver une population source couvrant correctement la population ciblée par l'étude ? Le critère de jugement M étudié survient-il rapidement ou longtemps après une cause possible E étudiée ? Etc.). **Le schéma d'étude doit donc être adapté à l'objectif de l'étude.**

I – Les enquêtes observationnelles sur données individuelles

Une **population ouverte** est une population en **constant renouvellement**, avec des arrivées et des sorties, des naissances et des décès, des immigrations, des émigrations, des perdus de vue (PDV)⁶. Les enquêtes observationnelles sur données individuelles consistent à sélectionner dans une population généralement ouverte, un échantillon de sujets pour chacun desquels on souhaite collecter des données pour déterminer le statut du ou des critères de jugement (notés M), d'une ou plusieurs expositions éventuelles (notées E), et éventuellement d'un ou plusieurs cofacteurs (notés F) pouvant intervenir dans la relation entre M et E.

Il existe trois types d'enquête observationnelle sur données individuelles : **enquêtes de cohorte, enquêtes cas-témoins, enquêtes transversales.**

I.A - Enquête de cohorte

Une enquête de cohorte permet de **suivre dans le temps** un échantillon de sujets **après leur inclusion** afin :

- D'identifier le **nombre de cas M+** d'un ou plusieurs critères de jugement **binaires M**, survenant au cours du temps **dans un échantillon de sujets tous M-** à l'inclusion (ex. : mortalité dans les 5 ans chez des patients inclus à l'occasion d'un diagnostic de cancer du sein ; survenue ou non dans les 6 mois d'une hospitalisation pour Covid-19 chez des patients inclus à l'occasion d'une vaccination anti-Covid-19, etc.).
- De suivre des **évolutions du niveau** d'un ou plusieurs critères de jugement **quantitatifs M** dont le niveau **est connu à l'inclusion chez les sujets de l'échantillon** (ex. : évolution du score de

⁶ Une **population fermée**, au contraire, est une population **qui ne grandit pas au cours du temps** (pas d'entrée de nouveaux membres) et qui ne perd aucun de ses membres sauf par décès (aucune sortie de sujets vivants).

dépression dans les 6 mois après inclusion pour une tentative de suicide ; évolution de la charge virale VIH plasmatique dans l'année suivant l'inclusion pour diagnostic de l'infection VIH, etc.).

Elle peut répondre à des **objectifs** :

- **Descriptifs** : estimer la « fréquence » de survenue de **M+** dans la population cible⁷, mesurée par un taux d'incidence ou par un risque (cf. Partie I.D), comparer éventuellement ce taux d'incidence à un taux d'incidence estimé dans une ou plusieurs autres populations, définies par exemple par un critère géographique, ou comparer l'évolution temporelle de ce taux d'incidence estimé pour la même population à plusieurs périodes. Les comparaisons temporelles et spatiales de taux d'incidence sont dans la mesure du possible **standardisées** sur la distribution du sexe et de l'âge (si ces informations sont disponibles), car celle-ci peut varier dans le temps et l'espace et expliquer une partie des différences observées (cf. Partie I.G).
- **Analytiques** : étudier la relation entre une exposition **E** et le critère de jugement **M**, en comparant des mesures de fréquences de survenue de **M+** entre sujets **E+** et **E-** par des tests et des mesures d'association (cf. Partie I.E).

Il existe **2 grands types de modalité de recrutement** des sujets constituant l'inclusion des sujets au temps initial (T0) dans une enquête de cohorte :

- ✓ Étude de cohorte à **recrutement transversal** (cf. Fiche 3) :
 - **Nécessaire pour les objectifs descriptifs**, et fréquente pour les objectifs analytiques, surtout lorsqu'on s'intéresse à plusieurs facteurs d'expositions **E** potentiellement associés à un ou plusieurs critères de jugement **M**.
 - L'échantillon initial est constitué comme pour une enquête transversale (cf. Fiche 6) : la sélection des sujets est effectuée dans une population source, soit de manière **exhaustive** (si la population est petite), **soit le plus souvent par tirage au sort** de sujets (cf. Partie I.G pour les modalités de tirage au sort). Mais contrairement à une enquête transversale, **un suivi longitudinal est prévu après l'inclusion**.
 - **Les sujets déjà porteurs de l'événement M+** à l'inclusion (lorsque le critère de jugement est binaire) **sont exclus** de l'échantillon initial pour le suivi longitudinal postérieur (ou du sous-échantillon analysé dans les cohortes à objectifs multiples).
 - Cet échantillonnage vise à avoir la meilleure **REPRÉSENTATIVITÉ** possible de la population cible lorsqu'elle comporte un ou plusieurs objectifs descriptifs, ce qui implique de rechercher une population source qui couvre le mieux possible la population cible (cf. Fiche 7) : files actives ou nouveaux patients de structures médicales concernées, personnes se présentant **consécutivement** dans une ou plusieurs structures susceptibles de recevoir la population cible (structures médicales pour des patients, écoles ou universités pour des élèves ou des étudiants, etc.), fichiers contenant les données identifiantes pour contacter les personnes sélectionnées (fichiers administratifs, fichiers de numéros de téléphone ou d'adresse par exemple pour le recrutement de la population générale), fichiers contenant directement les données nécessaires pour les analyses et permettant en particulier de déterminer le statut des participants vis-à-vis de **M** et **E** à l'inclusion et leur évolution au cours du suivi postérieur (bases de données hospitalières, fichiers nationaux des consommations de soins, etc.).
- ✓ Étude de cohorte de **type « exposés/non exposés »** (cf. Fiche 4) : forme particulière de cohorte dans laquelle on sélectionne séparément des sujets exposés **E+** et non exposés **E-** à l'inclusion, généralement utile lorsque les exposés sont rares dans la population d'intérêt, ce qui permet

⁷ Plus rarement en épidémiologie et en santé publique, l'objectif peut être de décrire l'évolution du niveau quantitatif de **M** dans le temps et l'espace, quand le critère de jugement est quantitatif.

d'équilibrer les groupes à comparer (cf. ci-dessous). Par exemple, en recrutant dans une entreprise, séparément des employés exposés à des produits chimiques (E+) et des employés administratifs de cette même entreprise (E-).

Suivi postérieur à l'inclusion : une cohorte est caractérisée par le suivi actif ou passif des sujets (= **étude longitudinale**) depuis leur inclusion, pendant une durée en général déterminée à l'avance dans le protocole de l'étude. Le recueil des données postérieures à l'inclusion permet d'identifier :

- La **survenue éventuelle de l'événement M+** étudié (tous les participants étant M- à l'inclusion) lorsque le critère de jugement est binaire (M+/M-) avec sa **date de survenue**, ou les **évolutions de niveau du critère M** s'il s'agit d'un critère quantitatif dont on dispose du niveau de base à l'inclusion.
- La **date de sortie anticipée de l'étude (=censure)** en raison d'un décès (si le décès ne fait pas partie du critère de jugement lui-même), d'un déménagement ou de l'absence de nouvelles à partir d'une date antérieure à la date de fin du suivi prévu (perdu de vue).

Pour réaliser l'analyse statistique, il est nécessaire de fixer une **date de point** (fin de l'étude ou date intermédiaire) pour faire le bilan de ce qu'il s'est passé depuis l'inclusion (nombre d'événements, de perdus de vue, temps de suivi de chaque sujet, etc.). Les informations éventuellement déjà recueillies après cette date de point (si le suivi continue) ne doivent pas être prises en compte dans cette analyse.

L'inclusion peut s'échelonner dans le temps, sur une journée ou sur plusieurs années. En revanche, la durée de suivi n'est pas la même pour tous les sujets. En effet, la fin du suivi pour chaque individu survient à des dates variables :

- Pour un sujet **présentant l'événement M+** étudié au cours du suivi dans la cohorte, sa date de fin de suivi correspond à la **date de survenue de M+**.
- Pour un sujet **perdu de vue** avant la fin de la période de suivi dans l'étude (déménagement à l'étranger par exemple), sans avoir développé M+ au cours du suivi observé, sa date de fin de suivi correspond à la **date de dernières nouvelles**.
- Pour un sujet **décédé sans avoir développé M+** (si le décès n'est pas l'événement étudié lui-même, ou n'est pas la circonstance de découverte de cet événement M+), sa date de fin de suivi correspond à la **date de décès**.
- Pour un sujet **n'ayant pas développé les événements ci-dessus**, la fin de suivi est la date prévue de fin d'étude. **La durée de suivi prévue par le protocole peut être la même ou non pour tous.**

Lorsque le suivi d'un participant est **censuré**, cela signifie qu'on dispose pour celui-ci d'une information incomplète : on ne sait pas si et quand le sujet aurait développé M+ après sa date de censure mais on sait qu'il n'a pas vécu l'événement M+ de l'inclusion jusqu'à « *au moins* » sa date de censure/fin de suivi.

Exemple : On décide d'inclure tous les patients atteints d'un cancer du pancréas à l'occasion du diagnostic entre le 1^{er} janvier 2020 et le 31 décembre 2022, afin d'étudier le risque de décès dans les 5 ans après le diagnostic. Deux possibilités :

- (1) *La durée maximale de suivi pour chaque sujet est de 5 ans à partir de l'inclusion. La fin de l'étude aura lieu au plus tard le 31 décembre 2027 si le dernier sujet a été inclus le 31 décembre 2022. **Deux événements censurant sont possibles** : être perdu de vue, ou être décédé pour une cause autre que M+ avant 5 ans. Si le patient a développé M+ après cette date (s'il ne l'avait pas développé avant), on ne peut pas le savoir.*
- (2) *Le suivi s'arrêtera au plus tard le 31 décembre 2024 pour tous, quelle que soit sa date d'inclusion. La durée de suivi n'est donc pas la même pour tous, même pour les sujets non perdus de vue, ne développant pas M+, ou ne décédant pas d'une autre cause que*

M. La date de fin d'étude constitue donc aussi un événement censurant, car la durée d'observation possible de M+ n'est pas la même pour tous.

Caractère **rétrospectif, prospectif ou mixte** d'une enquête de cohorte :

- **Cohorte prospective (ou à recrutement prospectif)** : la date de début des inclusions est postérieure à la date de finalisation du protocole de l'étude.
 - La sélection des sujets éligibles peut se faire à partir d'un fichier (la file active des patients d'un hôpital par exemple), ou au fur et à mesure que les personnes se présentent dans une structure de soins pendant une période donnée : l'étude peut être proposée à tous les sujets éligibles qui se présentent consécutivement, ou à une partie d'entre eux (par exemple 1 patient sur 2 ou sur 3... selon les effectifs attendus).
 - Attention, même si la sélection des sujets éligibles a été faite en une fois par tirage au sort dans le fichier des patients, l'inclusion peut nécessiter de contacter individuellement ces patients pour obtenir leur accord de participation et recueillir des données : l'inclusion peut donc s'étaler sur une ou plusieurs semaines, voire mois (il ne s'agit pas ici de la période de suivi longitudinal, mais juste de la période de recrutement).
- **Cohorte rétrospective (= cohorte historique)** : la date de début des inclusions est antérieure à la date de finalisation du protocole de l'étude.
 - L'échantillon initial est donc reconstitué *a posteriori* tel qu'on aurait pu le constituer dans le passé en sélectionnant un groupe de sujets (tous M-) à un temps initial passé. Ceci n'est donc possible que lorsqu'on dispose d'un fichier de sujets identifiables pour une telle étude dont on est capable de déterminer le statut vis-à-vis de M à l'inclusion même pour les sujets qui ne sont plus joignables au moment de la finalisation du protocole.
 - L'historique du suivi de ces sujets entre l'inclusion et la date de point est donc reconstitué rétrospectivement. Ceci n'est possible que si on peut disposer du suivi individuel des sujets depuis l'inclusion jusqu'à la date de point (niveau d'E initial et au cours du suivi, état de santé aux dernières nouvelles avec les dates disponibles pour chaque situation : vivant encore M- à la date de point, vivant M+ avec date de survenue, décédé encore M- à la date du décès, perdu de vue encore M- à la date des dernières nouvelles).

Exemple : pour étudier le risque de survenue de pathologies musculosquelettiques (M+) dans une entreprise de déménagement, on dispose du fichier des salariés ayant travaillé dans cette entreprise au cours des vingt précédentes années, et comportant les données de médecine du travail permet de repérer i) ceux qui ont développé M+ avec la date de survenue de M+, ii) les sujets décédés d'une autre cause que M+ avec la date de décès, iii) et les sujets qui ont quitté l'entreprise sans avoir développé M+ avec la date de départ et pour lesquels on n'a plus de nouvelles.
- **Cohorte mixte** : association des deux modes d'inclusion précédents (inclusion et suivi en partie rétrospectifs à partir d'une date spécifiée antérieure à la date de finalisation du protocole, puis poursuite des inclusions et/ou du suivi de manière prospective après cette date de finalisation).

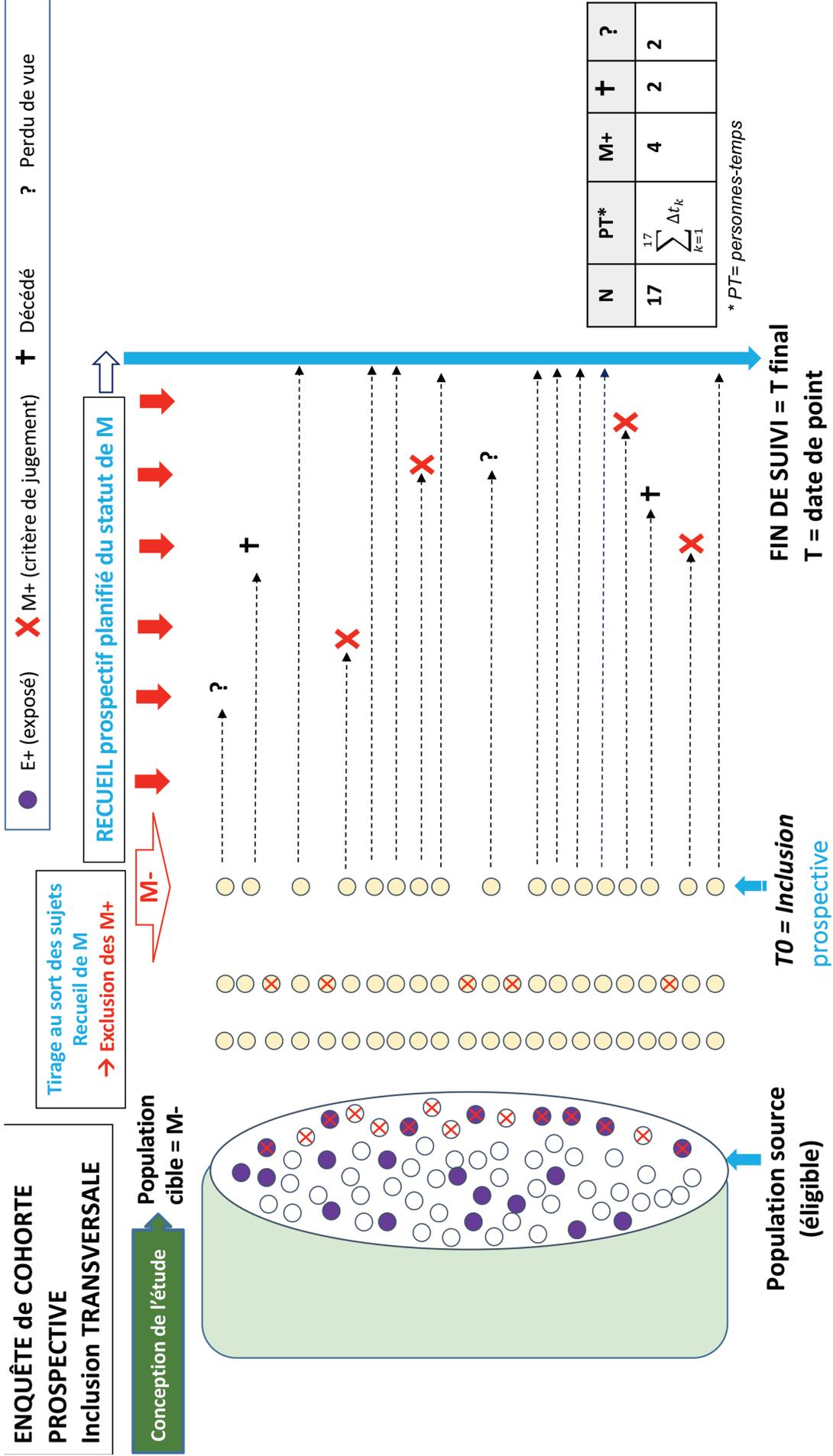
Les fiches 3 et 4 synthétisent les principales caractéristiques des enquêtes de cohortes à recrutement transversal et de type exposés / non exposés.

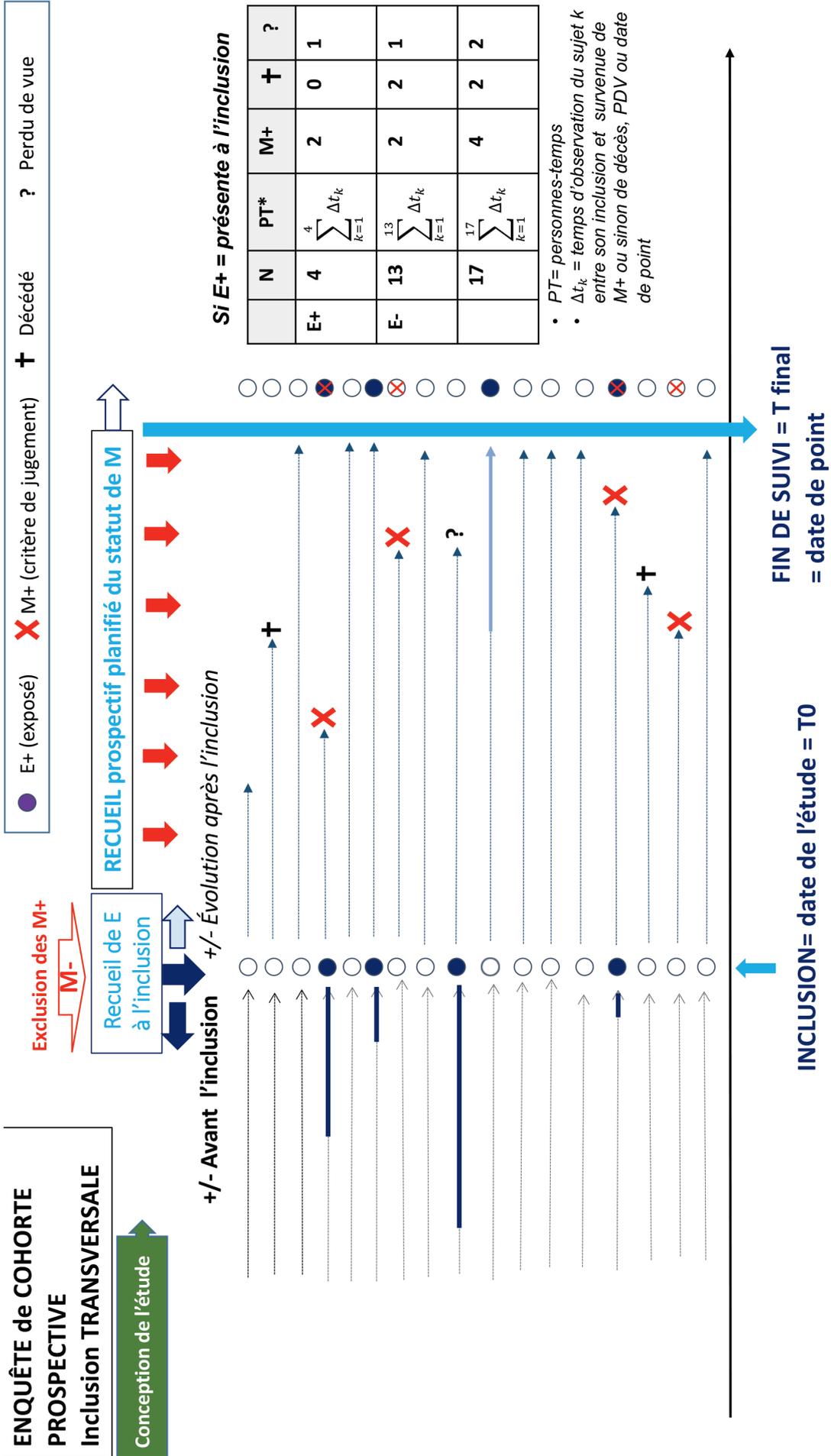


ENQUÊTE DE COHORTE RECRUTEMENT TRANSVERSAL

Objectif	Etude descriptive	Estimation de taux d'incidence de pathologie, comportements, etc., Evolution longitudinale d'un paramètre quantitatif au cours du temps, Etude pronostique : évolution d'une maladie (M = complications, récurrences, etc.).		
	Etude analytique (observationnelle)	Etude de l'association entre un critère de jugement M et un ou plusieurs facteurs de risque potentiel E (études étiologiques, pronostiques)		
	Etude (quasi) expérimentale	Evaluation de l'effet sur M d'une intervention E attribuée en situation quasi-expérimentale à tous les sujets (avec leur accord) au cours du suivi par exemple		
	Etude diagnostique	NON (sauf si un suivi longitudinal est nécessaire pour déterminer <i>a posteriori</i> le statut M+/M- au moment où le nouveau test à évaluer T a été réalisé)		
Variables clefs	M = critère de jugement E = exposition	Un ou plusieurs critères principaux ± critères secondaires, quantitatifs ou qualitatifs (situation classique : binaire M+/M-) Une ou plusieurs expositions selon l'objectif		
Sélection des sujets	Population cible S = Population source	Population non porteuse de M+ (si critère qualitatif binaire) à l'inclusion et susceptible de le devenir sur la période de suivi considérée Critères d'éligibilité spécifiques pour le recrutement en pratique : population générale ou spécifique (exemple population de patients pour des études pronostiques avec M = complication, récurrence, etc.) / Doit être le plus représentatif possible de la population cible si objectifs descriptifs ++		
	Modalités de sélection	<table border="1"> <tr> <td>Cohorte prospective Date de finalisation du protocole antérieure à la période d'inclusion Tirage au sort d'un échantillon dans la population source : - à partir d'un fichier initial, - ou par inclusion consécutive, de tout ou partie (1 sujet sur 2 ou 3, etc.) des sujets qui se présentent sur une période donnée Exclusion des sujets déjà M+ à l'inclusion (au fur et à mesure ou après sélection) Suivi longitudinal des sujets M-</td> <td>Cohorte rétrospective Date de finalisation du protocole postérieure à la période d'inclusion → une partie de sujets a déjà développé M+, ou sont décédés ou perdus de vue Tirage au sort d'un échantillon dans une population source présente au moment de la période d'inclusion visée, antérieure à la date de finalisation du protocole si liste disponible de ces sujets → reconstitution rétrospective de l'échantillon Exclusion des sujets déjà M+ à leur inclusion Reconstitution rétrospective du statut de E et M au cours du suivi : possible si base de données exhaustive préexistante (médico-administrative, liste d'employés, file active de patients, etc.)</td> </tr> </table>	Cohorte prospective Date de finalisation du protocole antérieure à la période d'inclusion Tirage au sort d'un échantillon dans la population source : - à partir d'un fichier initial, - ou par inclusion consécutive, de tout ou partie (1 sujet sur 2 ou 3, etc.) des sujets qui se présentent sur une période donnée Exclusion des sujets déjà M+ à l'inclusion (au fur et à mesure ou après sélection) Suivi longitudinal des sujets M-	Cohorte rétrospective Date de finalisation du protocole postérieure à la période d'inclusion → une partie de sujets a déjà développé M+, ou sont décédés ou perdus de vue Tirage au sort d'un échantillon dans une population source présente au moment de la période d'inclusion visée, antérieure à la date de finalisation du protocole si liste disponible de ces sujets → reconstitution rétrospective de l'échantillon Exclusion des sujets déjà M+ à leur inclusion Reconstitution rétrospective du statut de E et M au cours du suivi : possible si base de données exhaustive préexistante (médico-administrative, liste d'employés, file active de patients, etc.)
Cohorte prospective Date de finalisation du protocole antérieure à la période d'inclusion Tirage au sort d'un échantillon dans la population source : - à partir d'un fichier initial, - ou par inclusion consécutive, de tout ou partie (1 sujet sur 2 ou 3, etc.) des sujets qui se présentent sur une période donnée Exclusion des sujets déjà M+ à l'inclusion (au fur et à mesure ou après sélection) Suivi longitudinal des sujets M-	Cohorte rétrospective Date de finalisation du protocole postérieure à la période d'inclusion → une partie de sujets a déjà développé M+, ou sont décédés ou perdus de vue Tirage au sort d'un échantillon dans une population source présente au moment de la période d'inclusion visée, antérieure à la date de finalisation du protocole si liste disponible de ces sujets → reconstitution rétrospective de l'échantillon Exclusion des sujets déjà M+ à leur inclusion Reconstitution rétrospective du statut de E et M au cours du suivi : possible si base de données exhaustive préexistante (médico-administrative, liste d'employés, file active de patients, etc.)			
	Effectif Période d'inclusion (dates)	Déterminé pour l'ensemble de l'échantillon (pas de ratio fixé pour E+/E-) - Cohorte fixe (ou fermée) : inclusion des sujets éligibles à la même date ou période - Cohorte dynamique (ou ouverte) : inclusion des sujets éligibles au fur et à mesure qu'ils se présentent avec les critères d'éligibilité, jusqu'à la fin de l'étude (date d'entrée variable d'un sujet à l'autre).		
Recueil de données	Suivi longitudinal Recueil du statut de M Recueil de E (± F = facteurs de confusion)	TOUJOURS : suivi individuel entre inclusion (T ₀) et date de fin de suivi (T _f = date survenue de M+, du décès, de dernières nouvelles, date de point = fin d'étude) Recueil individuel de son évolution au cours du suivi indispensable Recueil si objectif(s) analytique(s) : statut à l'inclusion ± historique ± évolution après l'inclusion (= variables dépendantes du temps),		
Date de conception de l'étude		Antérieure à la période d'inclusion si cohorte prospective Postérieure à la période d'inclusion si cohorte rétrospective		

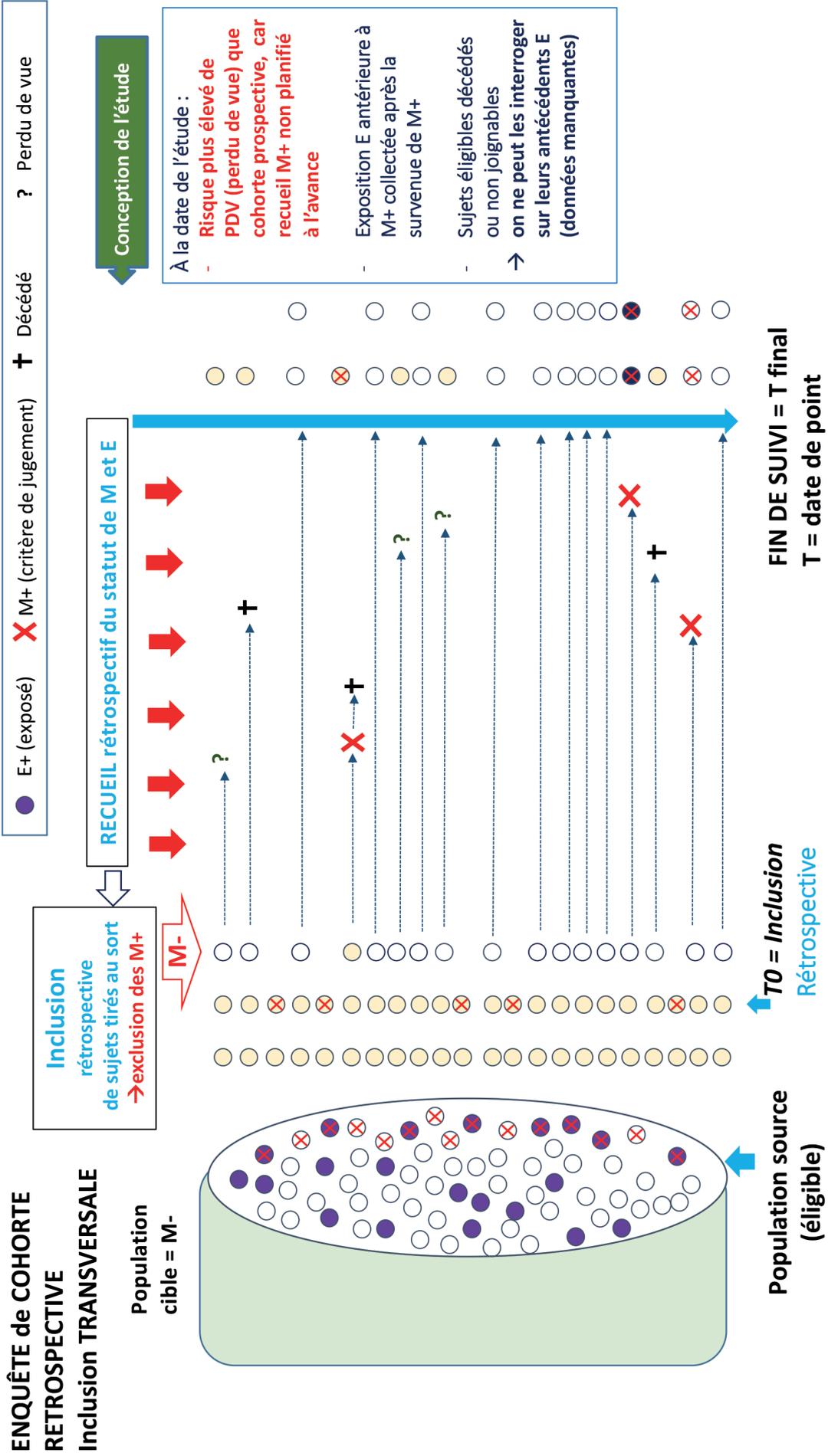
Fiche 3 • Enquête de cohorte recrutement transversal





Fiche 3 • Enquête de cohorte recrutement transversal

Fiche 3 • Enquête de cohorte recrutement transversal



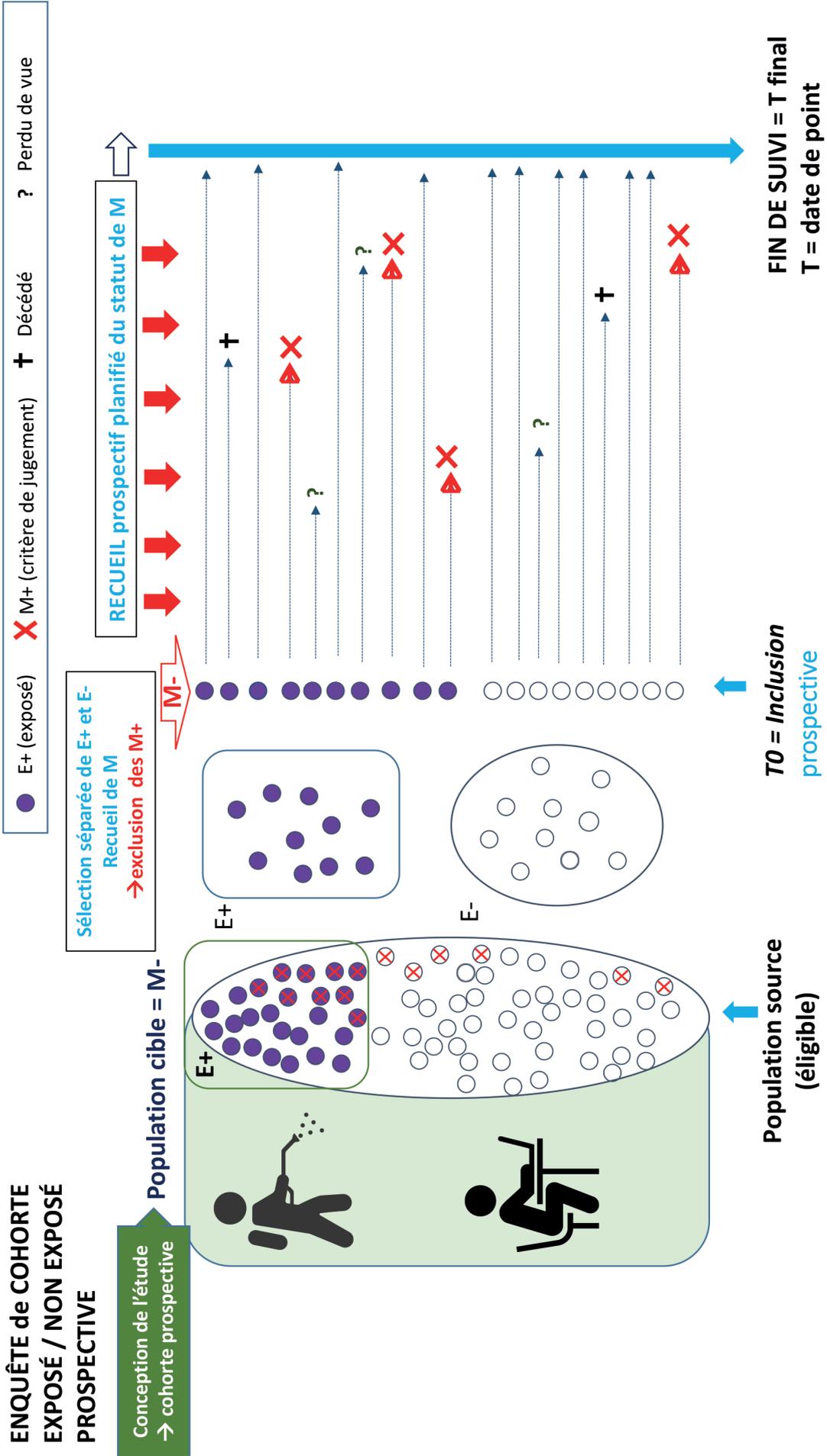


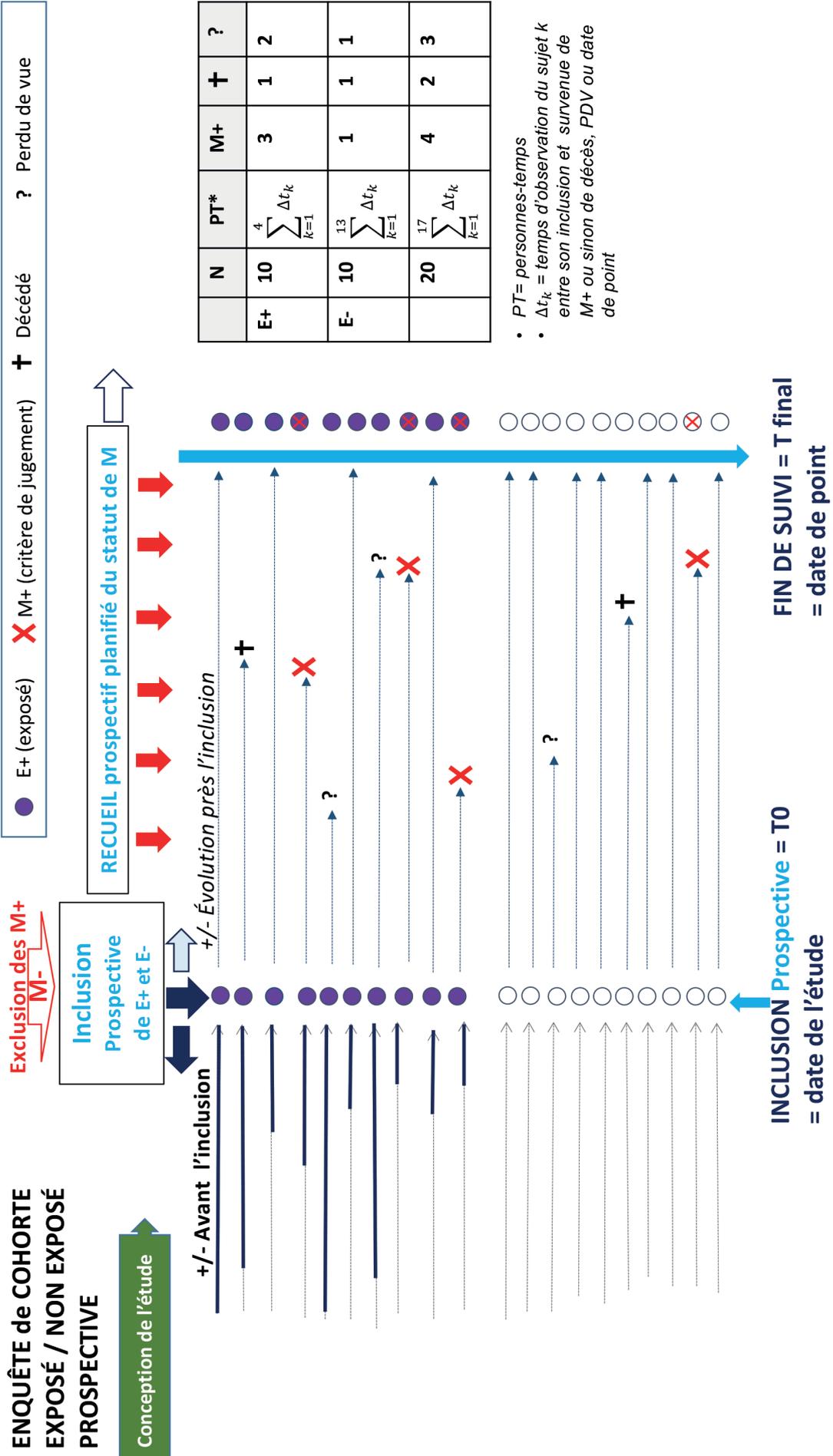
ENQUÊTE DE COHORTE DE TYPE EXPOSÉ/NON-EXPOSÉ

Objectif	Etude descriptive	NON							
	Etude analytique (observationnelle)	Etude de l'association entre un critère de jugement M et un ou plusieurs facteurs de risque potentiels E (étiologiques, pronostiques, etc.)							
	Etude (quasi)expérimentale	Evaluation de l'effet sur la survenue de M d'une intervention E attribuée en situation quasi-expérimentale à une partie des sujets sans randomisation							
	Etude diagnostique	NON							
Variables clefs	M = critère de jugement	Un ou plusieurs critères principaux ± critères secondaires, quantitatifs ou qualitatifs							
	E = exposition	Une seule exposition (sur laquelle se fait l'échantillonnage)							
Sélection des sujets	Population cible	Population non porteuse de M+ (si critère qualitatif binaire) à l'inclusion et susceptible de le devenir sur la période de suivi considérée							
	S = Population source	<p>1) Etude observationnelle : sélection séparée de sujets E+ et E- issus de la même population à la même période, tous M-, par exemple :</p> <table border="1" style="width: 100%;"> <thead> <tr> <th style="text-align: left;">Exposés E+</th> <th style="text-align: left;">Non exposés E-</th> </tr> </thead> <tbody> <tr> <td>○ Employé manipulant un produit chimique</td> <td>○ Employés non exposés de la même entreprise</td> </tr> <tr> <td>○ Sujets ayant reçu un traitement rare dans une structure médicale</td> <td>○ Sujets non exposés à ce traitement pour une indication similaire dans la même structure</td> </tr> </tbody> </table> <p>⚠ <i>Appariement et/ou stratification sur d'autres facteurs que E possible</i></p> <p>2) Etude quasi-expérimentale : sélection de sujets E- /M- qui pourront recevoir ou non E+ (à l'inclusion ou au cours du suivi, selon le design choisi).</p>	Exposés E+	Non exposés E-	○ Employé manipulant un produit chimique	○ Employés non exposés de la même entreprise	○ Sujets ayant reçu un traitement rare dans une structure médicale	○ Sujets non exposés à ce traitement pour une indication similaire dans la même structure	
	Exposés E+	Non exposés E-							
	○ Employé manipulant un produit chimique	○ Employés non exposés de la même entreprise							
	○ Sujets ayant reçu un traitement rare dans une structure médicale	○ Sujets non exposés à ce traitement pour une indication similaire dans la même structure							
	Modalités de sélection	<table border="1" style="width: 100%;"> <thead> <tr> <th style="text-align: left;">Cohorte prospective</th> <th style="text-align: left;">Cohorte rétrospective</th> </tr> </thead> <tbody> <tr> <td>Date de conception du protocole de l'étude antérieure à la période d'inclusion</td> <td>Date de conception du protocole est postérieure à la période d'inclusion → une partie des sujets a déjà développé M+, sont décédés ou perdus de vue</td> </tr> <tr> <td>1) Etude observationnelle : sélection séparée E+ et E-</td> <td>Etude observationnelle uniquement : possible si listes disponibles de sujets E+ et E- séparées → reconstitution rétrospective de l'échantillon</td> </tr> <tr> <td>2) Etude quasi-expérimentale : attribution raisonnée de E+ à tout ou partie des sujets</td> <td></td> </tr> </tbody> </table>	Cohorte prospective	Cohorte rétrospective	Date de conception du protocole de l'étude antérieure à la période d'inclusion	Date de conception du protocole est postérieure à la période d'inclusion → une partie des sujets a déjà développé M+, sont décédés ou perdus de vue	1) Etude observationnelle : sélection séparée E+ et E-	Etude observationnelle uniquement : possible si listes disponibles de sujets E+ et E- séparées → reconstitution rétrospective de l'échantillon	2) Etude quasi-expérimentale : attribution raisonnée de E+ à tout ou partie des sujets
Cohorte prospective	Cohorte rétrospective								
Date de conception du protocole de l'étude antérieure à la période d'inclusion	Date de conception du protocole est postérieure à la période d'inclusion → une partie des sujets a déjà développé M+, sont décédés ou perdus de vue								
1) Etude observationnelle : sélection séparée E+ et E-	Etude observationnelle uniquement : possible si listes disponibles de sujets E+ et E- séparées → reconstitution rétrospective de l'échantillon								
2) Etude quasi-expérimentale : attribution raisonnée de E+ à tout ou partie des sujets									
Effectif	Le ratio E+/E- est fixé pour l'étude de manière à surreprésenter les sujets E+ (car études menées en général pour des expositions rares)								
Période d'inclusion (dates)	<p>Cohorte fixe (ou fermée) : inclusion des sujets éligibles à la même date ou période (le plus souvent dans les cohortes E+/E-)</p> <p>Cohorte dynamique (ou ouverte) : inclusion des sujets éligibles au fur et à mesure qu'ils se présentent avec les critères d'éligibilité en parallèle du suivi des sujets déjà inclus éventuellement jusqu'à la fin de l'étude ; date d'entrée variable d'un sujet à l'autre</p>								
Recueil des données	Suivi longitudinal	TOUJOURS : suivi individuel entre inclusion (T ₀) et date de fin de suivi (T _f = date de survenue de M+, du décès, de dernières nouvelles, date de point = fin d'étude) <i>NB : reconstitution rétrospective impossible si étude quasi-expérimentale</i>							
	Recueil de M	Recueil de son évolution au cours du suivi individuel indispensable							
	Recueil de E (± F = facteurs de confusion)	Statut de E = modalité d'inclusion +++ / L'évolution du statut de E peut être recueillie au cours du suivi / Statut de F à l'inclusion ± historique ± évolution							
Date de conception du protocole	Antérieure à la période d'inclusion si cohorte prospective Postérieure à la période d'inclusion si cohorte rétrospective								

(1) Une cohorte exposé/non exposé prospective avec attribution aléatoire de l'exposition est appelé « essai randomisé » et non cohorte (par usage)

Fiche 4 • Enquête de cohorte de type exposé/non-exposé





Fiche 4 • Enquête de cohorte de type exposé/non-exposé