

---

# Épidémiologie

## Principes et méthodes quantitatives

J. Bouyer, D. Hémon, S. Cordier,  
F. Derriennic, I. Stücker, B. Stengel, J. Clavel

---



34376

**Inserm**



Institut national  
de la santé et de la recherche médicale

MD954/2

34376

(1)

# Épidémiologie

Principes et  
méthodes quantitatives



34376



J. Bouyer, D. Hémon, S. Cordier,  
F. Derriennic, I. Stücker, B. Stengel, J. Clavel

LES EDITIONS  
INSERM

# Sommaire

## INTRODUCTION

<b>1 – Contributions de l'épidémiologie à la recherche étiologique</b> .....	3
Études portant sur des groupes : variations d'incidence, corrélation entre les variations d'indicateurs collectifs d'exposition et de santé .....	7
Variations en fonction des caractéristiques générales des personnes .....	7
Variations de la fréquence des maladies en fonction du temps .....	9
Variations géographiques de la fréquence des maladies .....	11
Corrélations entre les variations temporelles d'indicateurs collectifs d'exposition et de santé .....	12
Corrélations entre les variations géographiques d'indicateurs collectifs d'exposition et de santé .....	13
Corrélations au niveau de groupes : intérêt et limitations .....	15
Études portant sur des individus : analyse de l'association entre facteur de risque et état de santé .....	17
Existence d'une association entre facteur de risque et maladie .....	18
Force de l'association entre facteur de risque et maladie .....	19
Prise en compte des autres facteurs de risque connus .....	20
Relation entre le degré d'exposition et l'importance de l'élévation du risque .....	21
Mise en évidence d'une hétérogénéité de l'exposition .....	23
Mise en évidence d'une hétérogénéité étiologique de la maladie .....	24
Exposition simultanée à plusieurs facteurs de risque .....	25
Références .....	27

## PARTIE I MESURES DE RISQUE ET MESURES D'ASSOCIATION

<b>2 – Mesures d'incidence et mesures du risque</b> .....	31
Introduction .....	32
Prévalence .....	32
Taux d'incidence .....	34
Risque de maladie .....	44
Incidence instantanée .....	46
Relations entre mesures du risque et mesures d'incidence .....	50

Notations .....	54
Références .....	55
<b>3 – Mesures d'association entre la maladie et un facteur de risque</b> .....	57
Modèle additif et modèle multiplicatif .....	58
Propriétés du risque relatif et de l'odds ratio .....	62
Mesure d'impact potentiel .....	69
Deux facteurs de risque ou plus .....	76
Annexe .....	84
Notations .....	85
Références .....	86
PARTIE II BIAIS	
<b>4 – Les sources de variabilité et d'erreurs en épidémiologie – Fluctuations d'échantillonnage et biais</b> .....	89
Population cible, population source .....	90
Fluctuations d'échantillonnage .....	92
Biais .....	94
<b>5 – Biais de sélection</b> .....	97
Définition du biais de sélection .....	98
Biais dans la constitution de l'échantillon .....	101
Biais dans la constitution du groupe de référence .....	106
Biais dans le suivi des groupes d'étude .....	113
Notations .....	114
Références .....	114
<b>6 – Biais de classement</b> .....	117
Définitions .....	118
Erreurs différentielles .....	119
Erreurs non différentielles .....	121
Erreurs de classement sur un tiers facteur .....	126
Minimiser les erreurs de classement .....	130
Notations .....	132
Références .....	132

<b>7 – Biais dus aux facteurs de confusion</b> .....	133
Biais de confusion dans une population .....	134
Le phénomène de confusion .....	134
Biais de confusion et autres types de biais .....	138
Rapport de confusion .....	139
Prise en compte d'un facteur de confusion dans une enquête épidémiologique .....	141
Définition .....	141
Méthodes de prise en compte d'un facteur de confusion au moment de la planification .....	142
Méthodes de prise en compte d'un facteur de confusion au moment de l'analyse (en l'absence d'interaction) .....	145
Quelles variables prendre en compte comme facteurs de confusion ? .....	148
Facteurs de confusion et interprétation causale .....	153
Annexe : Calcul du rapport de confusion - Conditions pour qu'une variable F soit facteur de confusion .....	155
Notations .....	156
Références .....	156
PARTIE III MÉTHODES STATISTIQUES USUELLES EN ÉPIDÉMIOLOGIE	
<b>8 – Notions de base en statistique</b> .....	161
Fluctuations d'échantillonnage .....	162
Intervalle de confiance .....	168
Tests comparatifs .....	169
Notations .....	178
Références .....	178
<b>9 – Méthodes d'estimation et de test</b> .....	179
Estimation .....	180
Estimation ponctuelle – Méthode du maximum de vraisemblance .....	180
Estimation par intervalle .....	183
Tests statistiques .....	188
Comparaison à une population de référence .....	189
Comparaison entre deux échantillons – Liaison entre deux variables .....	193
Échantillons appariés .....	201
Notations .....	205
Références .....	205
PARTIE IV PRISE EN COMPTE D'UN FACTEUR DE CONFUSION – ANALYSE MULTIVARIÉE	
<b>10 – Standardisation des mesures de risque et d'association</b> .....	209
Introduction .....	210
Standardisation directe .....	212

Standardisation indirecte .....	219
Standardisation directe versus standardisation indirecte .....	222
Notations .....	226
Références .....	226
<b>11 – Prise en compte au moment de l'analyse d'un facteur de confusion – Méthode d'ajustement de Mantel-Haenszel</b> .....	227
Introduction .....	228
Estimation du risque relatif et de l'odds ratio ajustés .....	230
Test de comparaison à 1 du risque relatif et de l'odds ratio ajustés .....	233
Test de l'interaction .....	236
Notations .....	238
Références .....	238
<b>12 – Analyse multivariée – Principes généraux</b> .....	239
Introduction .....	240
Les principaux modèles multivariés utilisés en épidémiologie .....	241
Estimation et tests des paramètres dans un modèle multivarié .....	244
Choix du modèle .....	252
Notations .....	265
Références .....	265
<b>13 – La régression logistique</b> .....	267
Définition .....	268
Interprétation des coefficients .....	270
Interaction dans la régression logistique .....	275
Estimation et tests des paramètres dans la régression logistique .....	277
Choix du modèle .....	281
Modèle logistique et type d'enquête .....	284
Annexe .....	288
Notations .....	289
Références .....	289
<b>14 – Le modèle de Cox</b> .....	291
Définition .....	292
Interprétation des coefficients .....	294
Hypothèse des risques proportionnels .....	296

Relation avec le modèle logistique .....	298
Estimation et test des coefficients .....	299
Notations .....	301
Références .....	301

## PARTIE V LES ENQUÊTES ÉPIDÉMIOLOGIQUES

<b>15 – Généralités sur les enquêtes épidémiologiques</b> .....	305
Introduction .....	306
Différents types d'enquêtes épidémiologiques .....	306
Enquêtes descriptives .....	309
Enquêtes étiologiques .....	313
Protocole d'une enquête étiologique .....	315
Stratégie d'analyse des enquêtes étiologiques .....	322
Références .....	329
<b>16 – Choix et évaluation des paramètres caractérisant l'état de santé</b> .....	331
Choix des paramètres caractérisant l'état de santé .....	332
Évaluation des paramètres caractérisant l'état de santé .....	334
Sources d'informations utilisables .....	340
Échelles de santé et variables composites .....	342
Références .....	346
<b>17 – Évaluation des expositions en épidémiologie – Exemple des expositions professionnelles</b> .....	349
Introduction .....	350
Les sources d'information sur les expositions professionnelles .....	350
Les méthodes de classification des expositions individuelles .....	352
Les sources d'erreur .....	360
Références .....	363
<b>18 – Enquêtes de cohorte</b> .....	365
Introduction .....	366
Échantillonnage .....	368
Le recueil d'information .....	372
Analyse .....	378

Conclusion .....	386
Notations .....	387
Références .....	387
<b>19 – Enquêtes cas-témoins</b> .....	389
Principe de base des enquêtes cas-témoins .....	390
Protocole .....	390
Analyse des résultats .....	400
Interprétation des résultats .....	401
Intérêt des enquêtes cas-témoins par rapport à d'autres approches .....	403
Références .....	404
<b>20 – Enquêtes transversales</b> .....	405
Introduction .....	406
Types d'enquête .....	406
Échantillonnage dans les enquêtes transversales .....	407
Mesures de l'exposition .....	413
Mesures de l'état de santé .....	416
Analyse statistique et interprétation .....	418
Conclusions .....	420
Références .....	420
<b>21 – Évaluation des actions de santé</b> .....	423
Introduction .....	424
L'essai randomisé .....	426
Expériences non randomisées .....	428
Les enquêtes d'observation .....	431
Principaux biais, critères de causalité .....	434
Références .....	437
 PARTIE VI LA PUISSANCE EN ÉPIDÉMIOLOGIE	
<b>22 – Calcul de la puissance, facteurs déterminant sa valeur</b> .....	441
Introduction .....	442
Expressions permettant de calculer la puissance .....	444
Les différents facteurs déterminant la valeur de la puissance, sources de perte de puissance .....	451

Notations .....	454
Références .....	455
<b>23 – Prise en compte de la puissance dans une enquête épidémiologique</b> .....	457
Puissance et protocole .....	458
Puissance et analyse statistique .....	467
Puissance et interprétation des résultats .....	472
Notations .....	477
Références .....	477
<b>Tables statistiques</b> .....	479
Table 1 : Table de la loi normale centrée réduite ( $z$ ) .....	480
Table 2 : Table de la loi de khi-2 ( $\chi^2$ ) .....	481
Table 3 : Table de la loi de Student ( $t$ ) .....	482
Table 4A : Valeur de $\phi$ en fonction de la puissance ( $1 - \beta$ ) pour $\alpha = 0,05$ . Test bilatéral .....	483
Table 4B : Valeur de $\phi$ en fonction de la puissance ( $1 - \beta$ ) pour $\alpha = 0,05$ . Test unilatéral .....	483
Table 5 : Valeurs de $\text{Arcsin } \sqrt{p}$ en fonction de $p$ .....	484
Table 6A : Nombre de sujets à inclure dans chaque groupe cas et témoins (respectivement exposés et non-exposés) ( $n_0 = n_1 = n$ ) pour garantir une puissance de 80 % en fonction de la fréquence d'exposition chez les témoins, $P_{E_0}$ (respectivement la fréquence de la maladie chez les non-exposés, $P_0$ ) et de la valeur de l'odds ratio (OR) .....	485
Table 6B : Valeur minimale de l'odds ratio garantissant une puissance supérieure ou égale à 80 % en fonction de la fréquence d'exposition chez les témoins, $P_{E_0}$ (respectivement la fréquence de la maladie chez les non-exposés, $P_0$ ), du nombre de cas et de témoins (respectivement exposés et non-exposés) ( $n_0 = n_1 = n$ ) .....	485
Table 7A : Valeur de la puissance en fonction du nombre de cas attendus ( $E$ ) et de la valeur du SMR .....	486
Table 7B : Valeur minimum du SMR que l'on peut mettre en évidence en fonction de la puissance souhaitée et du nombre de cas attendus ( $E$ ) .....	487
<b>Bibliographie</b> .....	489
<b>Index</b> .....	491

Ce livre est consacré aux principes fondamentaux de l'épidémiologie moderne. Il est plus particulièrement tourné vers l'épidémiologie analytique (ou étiologique) dont l'objectif principal est la recherche des facteurs de risque des maladies et la quantification de leur association avec le risque de maladie. Les six parties couvrent l'ensemble des questions qu'il est nécessaire de connaître pour mettre en œuvre et analyser une enquête épidémiologique.

• **Mesures de risque et mesures d'association** - Les mesures du risque (prévalence, taux d'incidence, incidence instantanée) et les mesures d'association (excès de risque, risque relatif, odds ratio, fraction de risque attribuable) utilisées en épidémiologie sont définies et comparées. Leur utilisation et leur interprétation sont discutées.

• **Biais** - Chacun des trois types de biais rencontrés en épidémiologie (sélection, classement, confusion) fait l'objet d'un chapitre. Une place importante est consacrée aux moyens d'éviter les biais dans la réalisation d'une enquête et aux possibilités de les corriger.

• **Méthodes statistiques usuelles en épidémiologie** - En rappelant les principales notions statistiques, cette partie permet de préciser les notations qui seront utilisées dans les chapitres suivants.

• **Prise en compte d'un facteur de confusion. Analyse multivariée** - La prise en compte de facteurs de confusion est l'un des objectifs principaux de l'analyse statistique des enquêtes épidémiologiques. Dans cette partie, les méthodes d'analyse univariées et multivariées sont exposées tant sur le plan de leurs principes généraux que sur celui de leur mise en œuvre pratique. Les auteurs se sont attachés à rendre ces méthodes abordables à des non-statisticiens sans pour autant masquer les difficultés qui leur sont liées. Une discussion très détaillée est consacrée à leurs stratégies d'utilisation et à l'interprétation de leurs résultats.

• **Les enquêtes épidémiologiques** - Le principe, le protocole et l'analyse des principaux types d'enquêtes épidémiologiques (cohorte, cas-témoins, transversale) sont exposés. De nombreux exemples illustrent les six chapitres de cette partie.

• **La puissance en épidémiologie** - Dans cette partie consacrée aux calculs de la puissance et aux moyens de la prendre en compte dans les enquêtes épidémiologiques, le propos dépasse nettement celui de la simple technique de calcul pour aborder les questions de stratégie d'enquête et d'analyse liées à l'amélioration de la puissance statistique.

Dans l'ensemble du livre, les auteurs ont eu le souci constant de ne pas séparer les considérations statistiques et épidémiologiques et de montrer les liens étroits qui les unissent.

L'expérience des auteurs dans les domaines de l'enseignement et de la recherche en épidémiologie a permis de donner à cet ouvrage un caractère pédagogique le rendant accessible à un large public, sans pour autant se limiter à un niveau d'exposé élémentaire.

