

Module I : « Apprentissage de l'exercice médical »

Sous module : « *La médecine fondée sur les preuves* »

QUANTIFICATION DU RISQUE

Dr Catherine Arnaud, Pr Sandrine Andrieu

I. GENERALITES : LA NOTION DE RISQUE

I.1 – Rappels sur les enquêtes étiologiques ou analytiques

Dans les enquêtes étiologiques, il s'agit toujours de comparer des groupes de sujets pour mettre en évidence l'association entre une exposition et une maladie.

- Les *enquêtes exposés - non exposés* permettent de comparer l'incidence d'une pathologie chez des sujets exposés à un risque à celle de sujets non exposés. Classiquement, les sujets sont inclus au moment de leur exposition et sont suivis au cours du temps pour recueillir l'information sur la survenue éventuelle de la maladie.
- Dans une *enquête cas-témoin*, les sujets sont inclus au moment de la survenue de la maladie, le recueil de l'information est rétrospectif et concerne l'exposition au facteur de risque. La comparaison porte donc sur la fréquence d'exposition au facteur de risque entre les cas et les témoins.

Les enquêtes étiologiques reposent donc sur la recherche et la quantification de l'association entre une exposition et une maladie. En considérant l'exposition et la maladie sous la forme de variables binaires (oui/non), on aboutit en fin d'enquête à la répartition suivante des N individus où a, b, c et d représentent les effectifs observés dans chaque groupe :

		Maladie		
		Oui	Non	
Exposition	Oui	a	b	$n_1 = a + b$
	Non	c	d	$n_0 = c + d$
		$m_1 = a + c$	$m_0 = b + d$	N

On peut donc calculer la fréquence de la maladie chez les sujets non exposés au risque ($f_0 = c/n_0$) et chez les sujets exposés au risque ($f_1 = a/n_1$). L'association entre exposition et maladie peut être étudiée en comparant les deux fréquences f_0 et f_1 par un test du χ^2 (recherche d'une **liaison statistique**). La deuxième étape va consister à savoir si l'association mise en évidence entre une exposition et une maladie correspond à une **relation de cause à effet**. La troisième étape cherche à **quantifier le risque** et fera l'objet du chapitre suivant (chap. 2).

1.2 - Risque et facteur de risque - Association statistique et causalité

La notion de causalité (cause nécessaire et suffisante) ne peut s'appliquer en épidémiologie de population en raison des difficultés d'interprétation liées à l'observation. Pour souligner la spécificité de leur approche, les épidémiologistes préfèrent utiliser le terme de **facteur de risque** à celui de cause. Un facteur de risque est une variable statistiquement liée à l'événement étudié et un « indicateur » du lien causal qui peut exister entre ce facteur et la maladie étudiée.

Une conclusion « causale » n'est donc pas accessible à partir d'une seule enquête d'observation. Mais dans ces situations, **différents arguments permettent de retenir le caractère causal d'une association** entre un facteur de risque et une maladie dans une expérience par ailleurs bien conduite :

- L'association entre exposition au facteur et maladie doit être retrouvée dans différentes enquêtes (populations et conditions différentes) : **argument de constance et de reproductibilité.**
- L'exposition au facteur a précédé la maladie : **argument de temporalité** et il existe une association entre le facteur et la maladie.
- La **force** (risque relatif¹ élevé) **et la spécificité** (à un facteur donné est associée une maladie donnée) **de l'association** (par ailleurs biologiquement plausible) est un argument de poids,
- de même, l'existence d'une **relation dose-effet** : quand la durée et/ou la dose d'exposition au facteur augmente, la fréquence de la maladie augmente.
- La suppression du facteur entraîne secondairement une diminution de la fréquence de la maladie.
- Enfin, la compatibilité et la cohérence avec d'autres données scientifiques est essentielle.

II. INDICATEURS DE RISQUE

II.1 - Excès de risque

On appelle excès de risque la différence :

$$RE = I_e - I_{ne}$$

Cette différence exprime **l'augmentation de l'incidence de la maladie chez les sujets exposés au facteur de risque par rapport à ceux qui ne le sont pas**. En effet, parmi les personnes exposées qui ont contracté la maladie, un certain nombre l'auraient contracté, même si elles n'avaient pas été exposées.

On ne peut pas obtenir l'excès de risque à partir des seuls résultats d'une enquête cas-témoins, faute de connaître l'incidence chez les non exposés.

II.2 - Risque relatif et odds ratio

On appelle **risque relatif (RR)** le rapport de l'incidence de la maladie chez les exposés sur l'incidence chez les non exposés. Il mesure l'augmentation du risque de maladie chez les sujets exposés par rapport aux sujets non exposés ; c'est une **mesure de l'intensité de l'association entre le facteur d'exposition et la maladie**.

¹ Voir paragraphe II.2

$$\text{RR} = \frac{I_e}{I_{ne}} = \frac{a}{n_1} / \frac{c}{n_0}$$

Le risque relatif s'interprète de la manière suivante :

- Un RR=1 signifie qu'il n'existe pas de différence d'incidence entre les sujets exposés et ceux qui ne le sont pas. Le facteur d'exposition étudié n'est pas associé à l'apparition de la maladie
- Un RR>1 signifie que le facteur est lié positivement à l'apparition de la maladie.
- Un RR<1 signifie que le facteur d'exposition est un facteur protecteur.

Le risque relatif peut être calculé dans une enquête prospective. Dans une enquête cas-témoins, l'incidence de la maladie n'est pas connue. On est donc amené à comparer la fréquence d'exposition chez les cas et les témoins. On définit donc un nouvel indicateur de risque, l'**odds ratio** (noté OR) qui est le rapport des cotes d'exposition chez les cas et les témoins.

Soient e_1 la fréquence d'exposition chez les malades : $e_1 = a / m_1$

et e_0 la fréquence d'exposition chez les témoins : $e_0 = b / m_0$

On appelle cote d'exposition le rapport $e / (1-e)$ soient $e_1 / (1-e_1)$ chez les malades et $e_0 / (1-e_0)$ chez les témoins.

$$\text{OR} = \frac{e_1}{1-e_1} / \frac{e_0}{1-e_0} = \frac{a/(a+c)}{c/(a+c)} / \frac{b/(b+d)}{d/(b+d)} = \frac{ad}{bc}$$

L'interprétation de l'odds ratio est identique à celle du risque relatif.

Dans le cas où la prévalence de la maladie est faible, l'odds ratio est une bonne approximation du risque relatif.

$$\text{RR} = \frac{a}{a+b} / \frac{c}{c+d} \approx \frac{ad}{bc} \quad \text{dans le cas d'une maladie rare car } a \text{ petit devant } b \text{ et } c \text{ petit devant } d$$

Le risque relatif ne dépend ni de la fréquence « spontanée » de la maladie, ni de la fréquence de l'exposition. De ce fait, il ne renseigne pas sur la proportion de cas liés au facteur de risque, ni parmi les exposés, ni dans la population générale. La notion du risque relatif est donc insuffisante pour juger de l'impact quantitatif d'un facteur de risque dans la population.

III. INDICATEURS D'IMPACT

III.1 - Fraction étiologique chez les exposés

La fraction étiologique FE est, par définition, **la proportion de cas attribuable à l'exposition parmi les sujets exposés.**

$$FE = \frac{I_e - I_{ne}}{I_e} = 1 - \frac{1}{RR} = \frac{RR - 1}{RR}$$

La fraction étiologique permet d'exprimer la relation entre l'exposition et la maladie par un indice qui fournit un renseignement important pour la prévention. En effet, pour un groupe de sujets exposés au facteur de risque, la fraction étiologique fournit directement la proportion de cas attribuable au facteur. Ainsi, même pour un risque relatif assez petit, par exemple RR=2, la FE=50%, ce qui signifie que sur 100 cas de la maladie survenus chez les exposés, 50 sont dus au facteur de risque. Autrement dit, la suppression du facteur de risque, si la relation entre exposition et maladie est causale, conduirait à une diminution de 50% du risque de maladie chez les sujets exposés.

La fraction étiologique du risque chez les exposés peut être calculée à partir d'une enquête de cohorte, et également à partir d'une enquête cas-témoins si la maladie est rare. Dans ce cas, $FR = \frac{OR - 1}{OR}$.

III.2 - Risque attribuable dans la population

C'est la proportion, parmi tous les cas dans la population cible, de ceux que l'on peut attribuer à l'exposition. Elle mesure l'impact global du facteur de risque dans la population en tenant compte de la proportion des personnes exposées.

Soit p_e est la prévalence de l'exposition dans la population (proportion de sujets exposés dans la population).

$$RA = \frac{\text{cas incidents attribuables}}{\text{cas incidents totaux}} = \frac{I_e p_e - I_{ne} p_e}{I_e p_e + I_{ne} (1 - p_e)} = \frac{p_e (I_e - I_{ne})}{I_{ne} + p_e (I_e - I_{ne})} = \frac{p_e (RR - 1)}{1 + p_e (RR - 1)}$$

$$RA = \frac{p_e (OR - 1)}{1 + p_e (OR - 1)}$$

selon le type d'enquête envisagé cohorte ou cas-témoins.

La fraction étiologique du risque dans la population est un paramètre fondamental en santé publique car elle permet de quantifier la proportion de cas évitables si on supprime l'exposition.

Par exemple, si 20% d'une population est exposée à un facteur de risque ($p_e=0,20$) et si le risque relatif correspondant à ce facteur vaut 2, alors $RA = \frac{0,2 (2 - 1)}{1 + 0,2 (2 - 1)} = 0,167$ ce qui signifie que parmi l'ensemble des cas observés dans cette

population, 16,7% sont attribuables à l'exposition au facteur de risque. Si on a observé 5 000 cas une certaine année, on peut dire que 833 parmi ceux-ci auraient pu être évités si le facteur de risque n'avait pas existé.

IV. INTERPRETATION DES RESULTATS

IV.1 – Biais et erreurs en épidémiologie : position du problème

Dans les études analytiques, la mesure de l'association entre un facteur de risque et une maladie se fait sur un échantillon de la population. Ainsi, la valeur observée de cette mesure (risque relatif, odds ratio) peut s'écarter plus ou moins de la réalité pour des raisons qui tiennent essentiellement aux fluctuations d'échantillonnage (erreurs aléatoires) ou aux biais (erreurs systématiques). La prise en compte des fluctuations d'échantillonnage intervient dans l'expression des résultats (on est amené à encadrer la valeur du risque observée par un intervalle de confiance) et dans l'interprétation du test statistique réalisé.

A l'inverse des erreurs aléatoires, le biais est une **erreur systématique** introduite entre les groupes à comparer (exposés / non exposés ; cas / témoins). Cette erreur peut affecter la planification et/ou le déroulement et/ou l'analyse de l'enquête. Qu'elle qu'en soit l'origine, elle peut conduire à une **estimation erronée de l'effet du facteur étudié sur l'incidence de la maladie**. Parfois, certains de ces biais peuvent être décelés a posteriori dans leur existence et leur grandeur ; dans ce cas, il existe des techniques **d'ajustement** des biais qui permettent d'en atténuer l'effet. Dans les enquêtes épidémiologiques, les biais sont le plus souvent difficiles à maîtriser et à quantifier. Il est donc important de chercher à en limiter les effets et leurs conséquences sur les conclusions que l'on peut tirer de l'enquête.

IV.2 - Biais qui affectent la constitution des groupes

Biais dans la définition initiale des groupes

1. Etudes de cohorte : c'est le fait d'être exposé ou non exposé qui définit les groupes.
 - Il y a peu d'ambiguïté dans les enquêtes prospectives vraies.
 - Dans les enquêtes prospectives historiques, la définition de l'exposition ou de la non exposition repose sur des sources d'information où les données n'ont pas été collectées dans ce but. En conséquence, il y a un risque de données manquantes ou imprécises soumises à des jugements interprétatifs.
2. Etudes cas-témoins : les groupes sont identifiés selon la présence ou l'absence d'une maladie. Il importe de vérifier que tous les cas sont effectivement atteints de la maladie (absence de "faux positifs") et que tous les témoins en sont indemnes (absence de "faux négatifs").

Représentativité des groupes

Il est fondamental que les groupes observés et comparés ne soient pas choisis de façon systématique parmi la population qu'ils sont censés représenter. On a affaire alors à des **BIAIS DE SELECTION** (ou d'échantillonnage). On regroupe sous ce terme tous les biais qui peuvent conduire à ce que les sujets effectivement observés dans l'enquête ne constituent pas un groupe représentatif des populations étudiées. Ce biais devient très gênant si les

deux groupes à comparer différent de façon systématique par un ou plusieurs caractères qui peuvent être liés au phénomène étudié.

Le **choix du groupe témoin** dans les enquêtes cas-témoins pose des problèmes de représentativité souvent difficiles à résoudre (recrutement en population générale, en structure hospitalière). Voici deux exemple de biais de sélection :

1) Dans les enquêtes sur les risques professionnels, on peut observer le "**healthy worker effect**" (effet du travailleur en bonne santé) lié au fait qu'en milieu de travail, les individus sont capables d'exercer une activité professionnelle régulière et sont donc a priori en meilleure santé que la population générale.

2) Un cas particulier de biais de sélection est celui pour lequel les sujets entrent ou non dans l'étude en fonction d'une décision qui leur est propre. On parle d'**auto sélection**, positive quand les sujets se portent volontaires pour participer à l'étude, négative quand les sujets refusent de répondre. Dans les deux cas, sauf preuve du contraire, il faut admettre que ces sujets sont différents de la population générale. Ils ont vis à vis du problème posé une attitude qui n'est pas neutre.

IV.3 - Les biais qui affectent la conduite de l'enquête

Biais de surveillance inégale

Le "danger" le plus grand est de ne pas être aussi rigoureux dans la surveillance et l'examen des deux groupes. Par exemple, dans une enquête exposés - non exposés, la recherche du diagnostic doit être aussi approfondie dans les deux groupes faute de quoi il peut y avoir un **biais de détection**.

Biais de mesure ou biais d'information

Les biais de mesure sont représentés par toutes les erreurs qui peuvent s'immiscer dans la mesure des phénomènes pris en compte : diagnostic ou exposition. Ils sont fréquents et plus difficiles à déceler dans les études cas-témoins. Ils peuvent avoir pour origine :

- **La subjectivité de l'enquêteur.** Ses préjugés, la façon de poser les questions peuvent induire les réponses "souhaitées", tant pour l'exposition que pour la maladie. Une façon de remédier à cette difficulté est, quand cela est possible, de faire travailler les enquêteurs à l'**aveugle**.
- **La subjectivité des enquêtés.** Elle peut apparaître sous forme de **biais de mémorisation**. Ce biais est spécifique des enquêtes cas-témoins. Très souvent, un malade se sentira plus concerné par l'enquête qu'un témoin et donnera donc des renseignements plus précis. Dans d'autres situations, le malade pourra avoir tendance à minimiser ou au contraire à exagérer l'exposition.
- **Les techniques de prise de mesures.** L'exactitude des mesures est fonction du matériel utilisé et des enquêteurs. Il est nécessaire d'avoir recours à un instrument de mesure (questionnaire, protocole d'interrogatoire) rigoureusement **standardisé** de façon à minimiser la part de l'interprétation. De même, les mesures instrumentales peuvent être soumises à des biais liés aux techniques et prises de mesures (exactitudes des mesures). Il peut s'agir d'une erreur «intrapersonnelle» (entraînement,

fatigue, stress), interpersonnelle (enquêteurs différents), technique (appareil ou réactifs de sensibilités différentes).

- **La codification et l'enregistrement des données.**

Les perdus du vue et non répondants

Ces difficultés existent essentiellement pour les études de cohorte. Elles sont d'autant plus probables que l'étude dure plus longtemps. Les causes de pertes de vue ne sont pas forcément également réparties dans les deux groupes (notamment une cause majeure peut être le décès lié à la maladie étudiée). Ce peut être aussi l'absence du domicile, l'incapacité de répondre pour cause de maladie ou de barrière linguistique, un déménagement, un refus ...

IV.4 - Les facteurs de confusion

Les facteurs de confusion sont liés à l'influence de tiers facteurs (liés à la fois à l'exposition et à la maladie) sur l'association entre exposition et maladie. Si par exemple, les sujets exposés sont plus âgés que les sujets non exposés, l'association observée dépend de l'effet de l'exposition mais également l'effet de l'âge.

Exemple : Le cancer du poumon est lié à la fois au tabagisme et à l'âge. Le risque de cancer augmente avec l'âge et plus l'âge est avancé, plus l'imprégnation tabagique a des chances d'être plus ancienne. Il faut donc neutraliser l'effet de l'âge avant de rechercher une causalité tabac/cancer du poumon.

Le contrôle des facteurs de confusion peut se faire :

- au moment de la planification de l'enquête : soit en choisissant pour chaque cas inclus un ou plusieurs témoin(s) identique(s) pour le facteur de confusion (méthode **d'appariement**), soit en réalisant une **stratification** sur le facteur de confusion (dans ce cas, la distribution du facteur de confusion est, par construction, comparable dans les deux groupes)
- au moment de l'analyse : il est possible de mettre en œuvre des méthodes d'analyses statistiques particulières (**méthodes d'ajustement**) qui permettent la prise en compte de telles situations.

IV.5 - Les biais liés à l'analyse des données

Outre les biais liés à la codification et à l'enregistrement des données, le choix des échelles peut influencer l'interprétation des

différences entre les groupes comparés. D'autres biais peuvent apparaître si on ne prend pas en compte les biais de sélection, si on modifie les hypothèses de départ ou les groupes de comparaison, si on modifie les erreurs de classification (risque α et β), si on explore de manière excessive toutes les associations possibles de cause à effet sans hypothèses de départ.