



L'épidémiologie pour tous

Erreurs

Rédigée par B. Toma, B. Dufour, J. Rivière et J.J. Bénet ; validée par le bureau de l'AEEMA

Une erreur correspond à une différence par rapport à la réalité. Elle peut être exprimée qualitativement (prétendre qu'un objet rouge est gris, condamner un innocent, etc.) ou quantitativement, sous forme d'une valeur numérique unique (l'estimation du poids d'un objet à 10 kg alors qu'il est de 7 kg) ou d'une fourchette de valeurs (l'estimation de l'âge d'un sujet entre trois et quatre ans alors qu'il est de six ans).

En épidémiologie, les erreurs et les risques d'erreurs sont fréquemment rencontrés.

Les erreurs en épidémiologie

Erreurs par excès et erreurs par défaut

Selon les cas, les erreurs peuvent appartenir à l'une ou l'autre de deux grandes catégories d'erreurs :

- Les erreurs « par excès » : il en est ainsi, pour une erreur exprimée quantitativement, lorsque le résultat mesuré ou estimé est supérieur à la valeur réelle (exemple ci-dessus de l'estimation du poids). Il peut également en être ainsi, pour une erreur exprimée qualitativement, en cas d'affirmation de présence de ce qui est absent (la condamnation

d'un innocent (considéré, à tort, comme coupable) ou lors de l'obtention d'un résultat positif à un test de dépistage alors que le sujet est indemne) ;

- Les erreurs « par défaut » : elles correspondent au type opposé d'erreur : l'exemple de l'estimation de l'âge pris ci-dessus (exprimée quantitativement) ou conclure à la bonne santé d'un malade par exemple, ne pas condamner un coupable ou obtenir un résultat négatif à un test alors que l'animal est infecté (expression qualitative).

Dans la vie de tous les jours et dans les activités quotidiennes de nombreuses professions (en particulier, mais pas seulement, en santé et en justice), chacun est exposé à ces deux types d'erreurs, aux conséquences différentes.

En statistique, d'autres expressions sont utilisées pour nommer ces deux types d'erreurs (cf. tableau 1) :

- L'erreur *par excès* est qualifiée *d'erreur de première espèce*, ou *erreur alpha* ;
- L'erreur *par défaut* est qualifiée *d'erreur de deuxième espèce*, ou *erreur bêta*.

En épidémiologie, on rencontre très souvent l'expression « risque d'erreur », sans pour autant que le type, alpha ou bêta, soit systématiquement mentionné.

Tableau 1 : Correspondance entre la réalité et la conclusion retenue

		<i>Réalité</i>	
		<i>OUI</i>	<i>NON</i>
Conclusion retenue	OUI	<i>Exact</i>	<i>Erreur de 1^{ère} espèce (alpha)</i>
	NON	<i>Erreur de 2^{ème} espèce (bêta)</i>	<i>Exact</i>

Lors d'un test statistique

Lors d'un *test statistique* (comme un χ^2 par exemple), ou lors d'expression d'une estimation sous forme d'un *intervalle de confiance* d'une valeur mesurée sur un échantillon, le risque d'erreur considéré (souvent de 5 % ou, parfois, de 1 %) est le **risque d'erreur alpha**, c'est-à-dire le risque de se tromper « par excès » : pour le test statistique, en affirmant l'existence d'une différence significative alors que l'écart observé n'est dû qu'au hasard et, pour l'intervalle de confiance, en pensant que la vraie valeur est comprise au sein de l'intervalle de confiance alors qu'elle ne s'y trouve pas.

Lors d'une démarche qualitative

Lors d'une *démarche qualitative* d'épidémiologie descriptive (savoir par exemple si une maladie est présente ou non dans une population), il faut tout d'abord choisir le pourcentage de prévalence que l'on veut être capable de détecter, c'est-à-dire le pourcentage pour lequel on veut avoir de grandes chances d'obtenir au moins une réponse positive au test de dépistage sur l'échantillon étudié. Il faut ensuite fixer un risque d'erreur toléré pour l'étude ; cette fois-ci, il s'agit du risque de ne pas obtenir au moins une réponse positive parmi l'échantillon, donc un risque d'erreur « par défaut » (**risque d'erreur bêta**), et donc de conclure à tort que la zone d'étude est indemne alors qu'elle ne l'est pas.

Avec un test de dépistage

On retrouve ces deux types d'erreurs pour les qualités intrinsèques d'un *test de dépistage* (ou de diagnostic), la sensibilité et la spécificité :

- Le complément de la sensibilité, c'est-à-dire la proportion de « faux négatifs », correspond à un risque d'erreur « par défaut » (proportion de sujets infectés ayant fourni une réponse faussement négative) ;
- Le complément de la spécificité, c'est-à-dire la proportion de « faux positifs », correspond à un risque d'erreur « par excès » (proportion de sujets indemnes ayant fourni une réponse faussement positive).

On peut noter que **la fréquence de ces deux risques liés à l'emploi de tests de dépistage est directement fonction de la prévalence de la maladie :**

- Pour le *risque d'erreur par défaut* : plus la maladie est fréquente dans la population, plus le nombre de faux négatifs est élevé ; d'où les difficultés pour identifier tous les animaux atteints dans un troupeau à pourcentage élevé de prévalence ;
- Pour le *risque d'erreur par excès* : plus la maladie est rare, plus le nombre de faux positifs est élevé ; d'où les difficultés rencontrées en fin de période d'éradication d'une maladie avec des troupeaux soupçonnés à tort d'être atteints.

A côté (et indépendamment) de ces deux catégories d'erreurs (« par excès » et « par défaut »), on peut distinguer deux autres types d'erreurs selon le caractère aléatoire ou systématique de leur survenue.

Les erreurs aléatoires et les erreurs systématiques

Les erreurs survenant *accidentellement*, notamment à cause d'une mauvaise manipulation sporadique de l'instrument de mesure, génératrices de résultats distribués de part et d'autre de la valeur véritable, sont les *erreurs vraies* ou *aléatoires*.



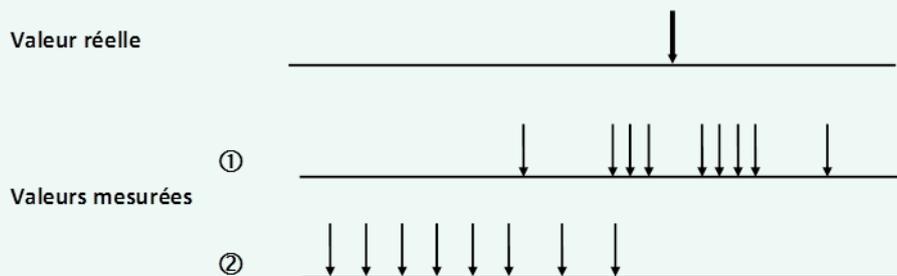
Les erreurs survenant **systematiquement**, en particulier en raison par exemple d'un mauvais réglage de l'appareil de mesure ou d'un choix orienté d'un échantillon dans une enquête sont appelées « **biais** ».

Leur distribution par rapport à la véritable valeur n'est pas aléatoire mais décalée systématiquement dans le même sens (à droite ou à gauche) par rapport à elle.

Ces deux types d'erreurs (aléatoires et biais) peuvent coexister (figure 1).

Figure 1 : Illustration de la différence entre erreur aléatoire et biais

- En ①, les valeurs mesurées comportent des différences (erreurs) par rapport à la valeur réelle. Ces erreurs sont aléatoires et conduisent à un résultat centré sur la valeur réelle.
- En ②, les valeurs mesurées comportent également des différences mais qui sont toutes orientées dans la même direction : les résultats sont biaisés car, à la fluctuation aléatoire des mesures, s'ajoute un caractère systématique qui entraîne un résultat faux, non centré sur la valeur réelle.



La gestion des erreurs

Il existe de nombreuses origines d'erreurs, notamment les instruments de mesure, les enquêteurs, les protocoles d'enquête, etc. Les solutions de gestion sont à adapter à chaque origine.

Par exemple :

- Pour les *instruments de mesure*, disposer d'informations sur la sensibilité et la spécificité du test, de façon à pouvoir estimer une prévalence

plus proche de la prévalence réelle à partir de la prévalence apparente ; pour des instruments utilisés sur le terrain, assurer la formation des utilisateurs afin de garantir la fiabilité de la mesure ([répétabilité](#) et [reproductibilité](#)) ;

- Pour les *enquêteurs*, avoir un protocole détaillé, précis et le respecter scrupuleusement ; effectuer la formation nécessaire à une standardisation de la collecte des données ;
- Pour le *protocole*, notamment tenter de limiter les biais d'échantillonnage.

Toma B. et al. : *Epidémiologie appliquée*, 2018, AEEMA éditeur, 614 p.

Terminologie : des liens hypertextes (en bleu souligné) vous permettent d'accéder aux définitions sur le site de l'AEEMA



Fiches AEEMA : [Fluctuations d'échantillonnage](#) ;

[Sensibilité \(individuelle ou collective\) d'un test de dépistage](#) ;

[Spécificité \(individuelle ou collective\) d'un test de dépistage](#).

[Valeurs prédictives d'un résultat](#).