

# **Statistiques et essais cliniques : concepts fondamentaux**

## **Introduction**

Les méthodes statistiques fournissent une comptabilisation formelle de l'origine de la variabilité des réponses à un traitement. Les statistiques permettent aux chercheurs cliniciens de tirer des conclusions raisonnables et exactes à partir des données recueillies et de prendre des décisions pertinentes en présence d'une incertitude. Les statistiques sont indispensables pour prévenir les erreurs et les biais en recherche médicale. Cet article couvre certains concepts statistiques fondamentaux et leur application aux essais cliniques.

## **Tests d'hypothèse**

Une hypothèse est une supposition ou un groupe de suppositions qui soit a) affirme quelque chose sur une base provisoire visant à orienter la recherche scientifique ou b) confirme quelque chose comme étant très probable à la lumière des faits établis.

Dans cet article, nous nous intéressons à la première forme d'hypothèse, par exemple l'affirmation qu'un nouveau traitement contre une maladie est meilleur que la norme de soins existante. Si le nouveau traitement est le traitement B et la norme de soins est le traitement A, alors l'hypothèse affirme que B est meilleur que A.

On pourrait penser que les chercheurs vont s'efforcer de prouver cette hypothèse, mais ce n'est pas le cas. En fait,

l'objectif visé est traité de façon indirecte. Au lieu d'essayer de prouver l'hypothèse B, la méthode scientifique suppose qu'en fait, A est vrai (qu'il n'y a pas de différence entre la norme de soins et le nouveau traitement). Cela s'appelle l'hypothèse nulle. Les chercheurs essaient ensuite de réfuter A, c'est-à-dire de démontrer que l'hypothèse nulle est fausse. S'ils y parviennent, s'ils démontrent que l'hypothèse A est fausse, et que la norme de soins n'est pas meilleure que le nouveau traitement, il en résulte que B est vrai et que le nouveau traitement est meilleur que la norme de soins.

## **Pourquoi procéder ainsi ?**

Il n'y a pas de réponse simple ; cette méthode largement utilisée est devenue la méthode scientifique moderne ; une analogie juridique permet peut-être d'y voir plus clair. L'hypothèse nulle correspond à la situation où la connaissance actuelle (si l'on fait une analogie avec un tribunal, c'est la situation de l'accusé qui est innocent), qu'il faut tenir pour vraie tant que rien ne prouve le contraire. Mais si nous voulions démontrer l'hypothèse alternative (l'inverse de l'hypothèse nulle), alors en fait, l'accusé est coupable.

Pour éclairer cette explication, on peut également citer Albert Einstein :

*« Aucune expérimentation ne pourra jamais démontrer que j'ai raison ; une seule expérience suffit à prouver que j'ai tort. »*

Ceci semble suggérer qu'il est plus rigoureux et plus faisable d'essayer de prouver que l'hypothèse nulle est fausse ou erronée que d'essayer de démontrer que l'hypothèse alternative est vraie. Notons que cela N'EXPLIQUE PAS vraiment pourquoi les chercheurs adoptent cette méthode, mais peut-être que cela peut nous aider à comprendre et à accepter plus facilement un concept difficile.

# Erreur de Type I et II

Le tableau ci-dessous permet de comprendre la différence entre l'erreur de Type I (faux positif) et l'erreur de Type II (faux négatif).

	<b>L'hypothèse nulle est vraie</b>	<b>L'hypothèse nulle est fausse</b>
L'hypothèse nulle est rejetée	Erreur de Type I « Faux positif »	Résultat correct « Vrai positif »
Echec du rejet de L'hypothèse nulle	Résultat correct « Vrai négatif »	Erreur de Type II « Faux négatif »

Cela reste très compliqué ; pour exprimer ce concept en termes plus simples, voici un exemple brut :

- L'erreur de Type I peut tuer un patient ; imaginez une étude qui conclut à tort que la norme de soins n'était pas meilleure que le nouveau traitement ; le nouveau traitement est administré aux malades et les résultats sont catastrophiques. L'erreur de Type I entraîne la détection d'un effet qui n'existe pas.
- L'erreur de Type II signifie que des recherches potentiellement utiles sont perdues. Ces recherches auraient peut-être pu s'avérer utiles, mais comme aucune recherche complémentaire n'est effectuée, les patients ne subissent pas de préjudice. L'erreur de Type II empêche de détecter un effet qui existe.

Il est donc clair que l'erreur de Type I est plus grave pour les patients que l'erreur de Type II.

## Niveau alpha

Le niveau alpha est la probabilité de commettre une erreur de Type I. Ce niveau est affecté par la taille de l'échantillon et par la puissance statistique du test.

# Puissance statistique

La puissance d'un test statistique est la probabilité qu'il engendre le rejet d'une hypothèse nulle ; en d'autres termes, c'est la capacité du test à détecter un effet, si cet effet existe. On peut également dire que la puissance d'un test est la probabilité de NE PAS commettre une erreur de Type II.

## Valeurs p

Les valeurs p, ou valeurs de probabilité, pondèrent la puissance d'une preuve sur une échelle de 0 à 1. Une valeur p faible (typiquement inférieure à 0,05, ou 5 %) indique qu'il y a une forte preuve *contre* l'hypothèse nulle, qui pourrait conduire à rejeter l'hypothèse nulle, tandis qu'une valeur p plus élevée (supérieure à 0,05) indique le contraire.

## Corrélation et causalité

Lors de l'analyse des résultats d'un essai, il est important de ne pas oublier que corrélation n'est pas causalité. On parle de corrélation lorsque deux variables sont en relation réciproque ; toutefois, cela n'implique pas qu'il s'agit d'une relation de causalité (il existe une association entre les deux variables). Voici un exemple faisant intervenir le traitement hormonal de substitution (THS) et les coronaropathies (CPT), les femmes sous THS ayant un risque moindre de CPT. Cela n'était cependant pas dû au THS lui-même, mais au fait que le groupe de patientes sous THS tendait à appartenir à une catégorie socio-économique plus élevée, bénéficiant d'une alimentation supérieure à la moyenne et pratiquant régulièrement un sport.

Une causalité peut être observée lorsqu'un facteur entraîne un résultat. Un facteur de causalité est souvent la cause partielle d'un résultat. Pour distinguer une corrélation d'une causalité, il est important d'enregistrer le maximum

d'informations sur les participants des essais. Il est également nécessaire de respecter rigoureusement la méthodologie scientifique dans la conception des essais cliniques et d'évaluer les biais possibles de l'essai.

## **Manipulation des données**

La manipulation de données consiste à rapporter sélectivement des données de manières erronées ou à créer de faux résultats. Par exemple, il y a manipulation si des données qui ne concordent pas avec les résultats escomptés sont intentionnellement supprimées afin d'augmenter la proportion de résultats qui confirmeraient l'hypothèse énoncée. Lorsqu'un chercheur élimine les valeurs extrêmes (résultats atypiques très éloignés de la masse des résultats), il doit vérifier qu'il s'agit réellement de valeurs extrêmes et non de résultats qui diffèrent simplement des résultats escomptés ou souhaités. Autre exemple : on parlerait également de manipulation si un logiciel de recueil de données générerait aléatoirement tout un jeu de données à partir du recueil des résultats d'un seul patient.

## **Transformation de données**

La transformation de données est l'application d'une formule mathématique à certaines données obtenues dans un essai. La transformation est souvent utilisée pour que la présentation des données soit plus claire ou plus facile à comprendre. Par exemple, si l'on mesure l'efficacité d'un carburant pour les voitures, il est naturel de mesurer cette efficacité sous forme de kilomètres par litre. Toutefois, s'il fallait évaluer le volume de carburant supplémentaire nécessaire pour accroître la distance parcourue, elle serait exprimée en litres par kilomètre. L'utilisation d'une formule incorrecte pour obtenir la nouvelle présentation de ces données dans ce cas affecterait les résultats globaux de l'essai.

# Fusion de données

La fusion consiste à associer des données issues de différentes études afin d'obtenir une meilleure compréhension de la situation. Une des formes les plus courantes est la méta-analyse, qui consiste à rassembler les résultats de différents essais publiés pour les agréger et les comparer. Il est important dans le cadre d'une méta-analyse, de vérifier soigneusement que les essais avaient des méthodologies identiques ou comparables. Toute différence dans la conception doit être prise en compte afin d'éviter les variables sous-jacentes différentes (variables confondantes). Par exemple, la fusion serait incorrecte si les données de plusieurs essais étaient regroupées alors qu'elles concernent différentes espèces de souris.

A2-4.33.3-V1.1