

C'est quoi cette taille d'effet ?

Mardi 6 Juin

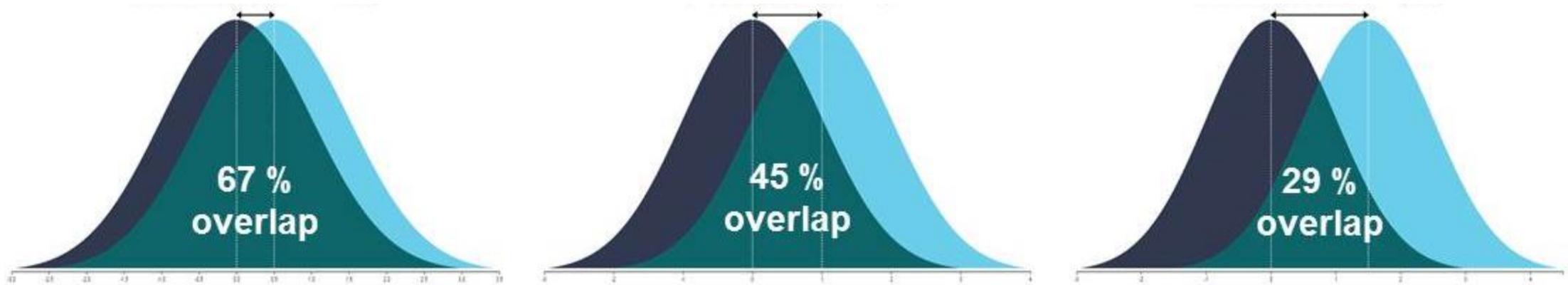
Orateurs : Pr. E. Tirelli et Pr. AFr. Donneau

En plus de la p-valeur, il est de plus en plus souvent demandé de documenter les résultats à l'aide de *taille d'effet* et de *puissance statistique*

C'est quoi cette taille d'effet ?

Définition de la taille d'effet (1/2)

La taille d'effet est une quantification standardisée d'une différence entre deux valeurs (e.g. moyennes, proportions) ou d'une association entre deux variables. Il s'agit d'un indice, score numérique, sans unité.



Source: modifiée de <https://eda.nc3rs.org.uk/experimental-design-group#variability>

Définition de la taille d'effet (2/2)

La famille « d » ou mesure des différences normalisées, l'amplitude d'une différence entre deux groupes indépendants ou dépendants (variable continue).

Exemples: d de Cohen, g de Hedges, psy de Steiger ...

La famille « r » ou degré d'association entre deux variables continues.

Exemples: r de Pearson, coefficient de détermination (R-carré), f-carré et f de Cohen, q de Cohen, éta-carré, éta-carré partiel, oméga-carré ...

La famille « OR » ou degré d'association entre variables catégorielles.

Exemples: rapport de cotes (OR) , risque relatif (RR), différence de risques, h de Cohen, coefficient phi, V de Cramer...

Comment obtenir la taille d'effet ?

- Calcul assez simple basé sur des données issues de l'application du test d'hypothèses :

- Pour la comparaison de deux populations indépendantes (*d* Cohen)

$$d_s = \frac{\bar{x}_1 - \bar{x}_2}{s_p} \text{ où } s_p \text{ est l'écart-type poolé.}$$

- Pour la comparaison de deux populations appariées (*d* Cohen)

$$d_z = \frac{\bar{x}_D}{s_D} \text{ où } \bar{x}_D \text{ la moyenne et } s_D \text{ l'écart-type des différences des données appariées.}$$

- Attention aux différents *d* Cohen !
- Pas toujours disponible par défaut dans les sorties des logiciels d'analyses – programmations supplémentaires. R: Package 'effectsize'

Interprétation d'une taille d'effet

- Chaque taille d'effet a sa propre interprétation (d Cohen, r, ...)

Effect size	<i>d</i>	Reference
Very small	0.01	Sawilowsky, 2009
Small	0.20	Cohen, 1988
Medium	0.50	Cohen, 1988
Large	0.80	Cohen, 1988
Very large	1.20	Sawilowsky, 2009
Huge	2.0	Sawilowsky, 2009

Effect size	<i>r</i>
Small	0.10
Medium	0.30
Large	0.50

- Une même taille d'effet peut être relevante dans un contexte scientifique mais pas du tout dans un autre, même dans le même domaine ou la même discipline!

Et par rapport à la p -valeur ? (1/2)

p -valeur = significativité statistique d'un test d'hypothèses

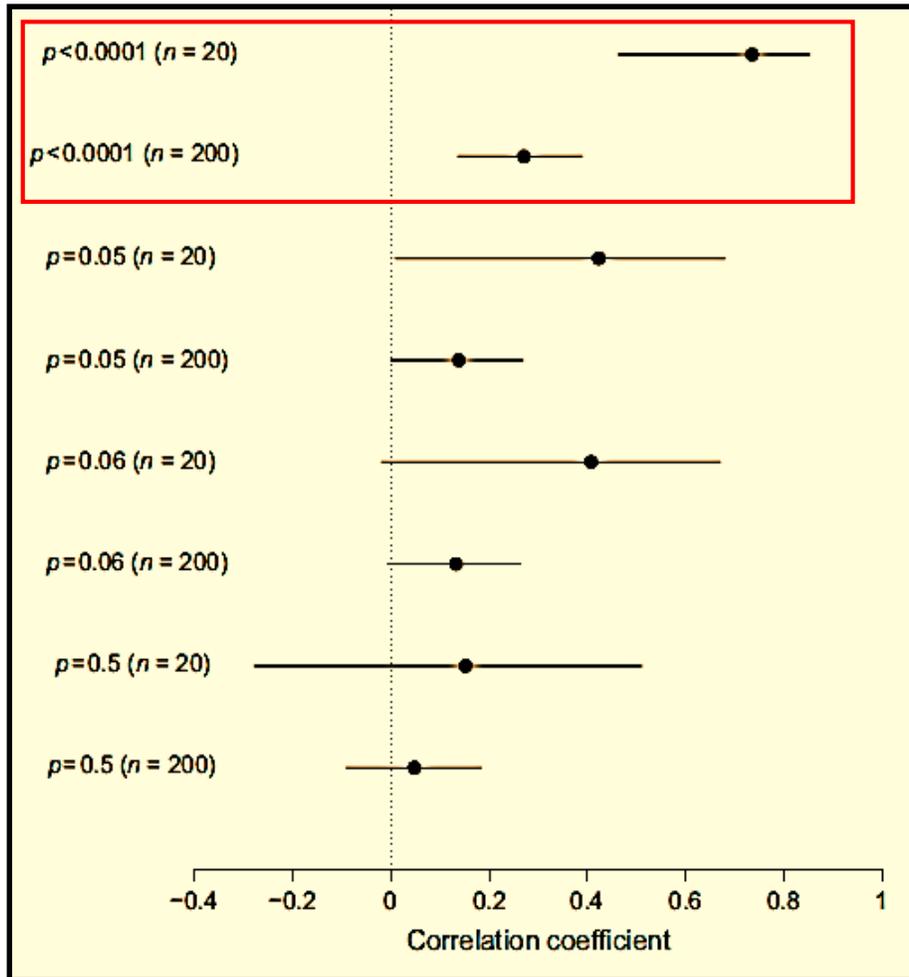
Taille d'effet = significativité scientifique du résultat du test d'hypothèses

Contrairement à la p -valeur, la taille d'effet est indépendante de la taille de l'échantillon (n) et de la puissance.

Des études réalisées sur de grands échantillons et présentant un petit effet peuvent facilement atteindre une signification statistique ($p < 0.05$).

A l'inverse, des études réalisées sur de petits échantillons, mais présentant une taille d'effet importante risquent de ne pas atteindre la signification statistique.

Et par rapport à la p -valeur ? (2/2)



Dans cet exemple, les deux p -valeurs « hautement significatives » présentent des tailles d'effet différentes.

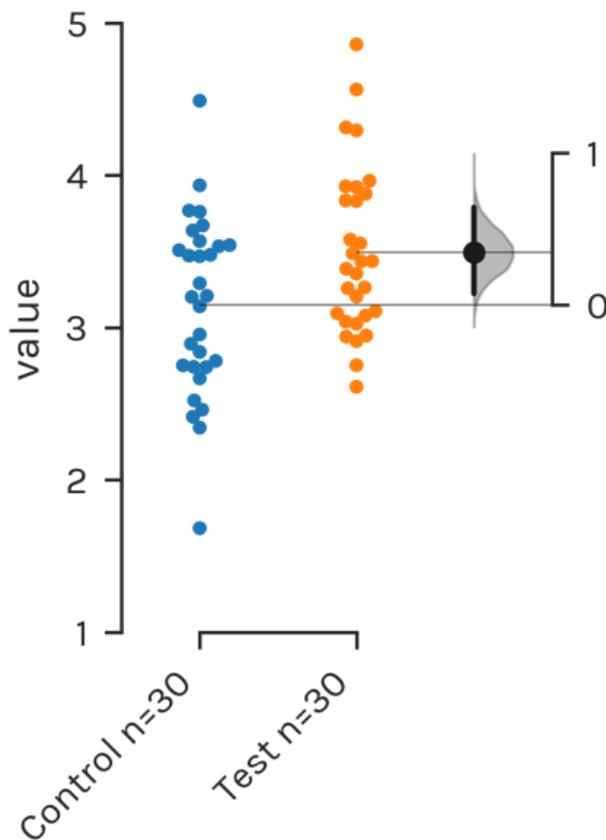
On observe des tailles d'effets différentes pour les trois autres p -valeurs ($p=0,05$, $p=0,06$ et $p=0,5$) avec des intervalles de confiance très différents.

Notez aussi que les comparaisons sont basées sur deux tailles d'échantillon différentes ($n=20$ ou $n=200$).

Source: Nakagawa & Cuthill (2007) *Biological Reviews* 82: 591-603.

Présentation des résultats (1/2)

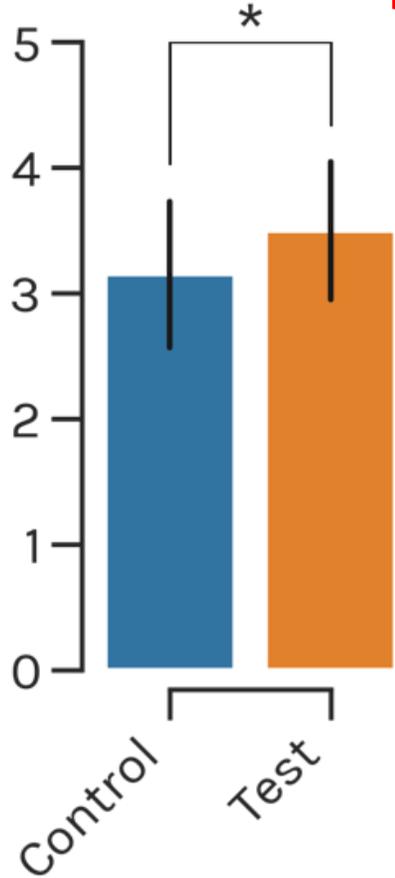
Exemple de représentation graphique transparente de résultats



- Toutes les valeurs obtenues et leur dispersion sont visibles (n=30).
- La taille de l'effet (différence non-standardisée) est représentée. La différence standardisée (d de Cohen) aurait pu aussi être représentée (échelle spécifique).
- La précision de l'estimation est donnée par l'intervalle de confiance à 95% (avec la distribution des tailles d'effets possibles en gris).
- L'usage direct de la p -valeur est ici moins indispensable et laisse la place à l'appréciation scientifique directe de l'effet via sa taille (importance factuelle).

Présentation des résultats (2/2)

Exemple de représentation graphique classique de résultats



- Les valeurs brutes ne sont pas visibles.
- Aucune indication sur la taille de l'effet.
- L'astérisque indique l'obtention d'un effet statistiquement significatif, ce qui n'informe aucunement sur l'effet en soi.
- Les barres verticales représentent ici un intervalle de confiance sur la moyenne.

Remarque: un nombre croissant mais encore modeste de revues scientifiques déconseillent - voire bannissent - ce type de représentation des résultats.

Rôle de la taille d'effet dans le calcul du « n » (1/2)

AVANT LA RECHERCHE

Taille d'effet hypothétique (ou prospective). A identifier sur la base de tailles d'effet plausibles et surtout justifiables par le contexte scientifique (signifiante).

&

Puissance prospective (*a priori*). Permet de calculer les effectifs indispensables à la détection par le test de la taille d'effet hypothétique (à une puissance élevée).

APRES LA RECHERCHE

Taille d'effet observée (et IC). A discuter en rapport avec la taille d'effet hypothétique, d'autres tailles d'effet possibles et surtout le contexte scientifique (signifiante).

&

La puissance observée (*a posteriori*). Son interprétation rétrospective est fallacieuse et n'ajoute rien au résultat statistique (car liée à la P-valeur). A bannir.

Rôle de la taille d'effet dans le calcul du « n » (2/2)

Divers domaines de recherche ont adopté des expressions spécifiques pour désigner la taille d'effet minimalement intéressante (la plus petite possible).

Celle-ci peut servir de taille d'effet hypothétique pour réaliser le calcul de taille d'échantillon - une taille d'effet minimalement détectable par le test statistique effectué avec une puissance minimale acceptable et à un risque alpha maximal justifiable.

MDE = « Minimally DETECTABLE Effect »

MDD = « Minimally DETECTABLE Difference »

Pour conclure...

L'interprétation des résultats basée uniquement sur la *p-value* fait perdre de vue la significativité scientifique des résultats ➡ Absence de discussion des effets observés (*taille d'effet*)

- Important de présenter la taille d'effet dans les articles/rapports,
- La taille d'effet permet d'objectiver vos résultats. Elle représente ce que vous recherchez depuis le début de votre étude!
- La taille d'effet est un gage de qualité de votre recherche.

Pour suivre nos activités ou prendre contact avec nous:



Formation d'une journée en présentielle
sur la taille d'effet prévue AA 23-24

LinkedIn: <https://www.linkedin.com/company/centre-b-stat/>