

## Contexte

Dans le domaine de la recherche clinique, l'essai randomisé contrôlé avec comparaison directe reste le *gold standard* pour évaluer l'efficacité d'un nouveau traitement ou dispositif médical. Cependant, de nouvelles approches méthodologiques comme les méthodes de **comparaison indirecte ajustée sur la population PAIC** (*Population Adjusted Indirect Comparison*), qui reposent sur des **données externes**, se développent. Parmi celles-ci, l'intégration d'un **bras de contrôle virtuel/synthétique simulé** peut apporter plusieurs avantages tant dans le cadre d'études interventionnelles qu'observationnelles. Ces bras virtuels simulés, construits à partir de données agrégées externes, offrent une alternative aux comparaisons directes traditionnelles.

Dans ce travail, nous nous appuyons sur une étude en comparaison directe, qui démontre la **non-infériorité** d'un nouveau traitement (N) utilisé contre la **sécheresse oculaire** par rapport à un traitement standard (S). Dans cette étude, la non-infériorité du traitement N est démontrée avec un seuil de non-infériorité fixé à  $\delta = -2$ . Nous utilisons les données agrégées d'une étude externe publiée ayant un bras traité avec S (voir Figure 1), données qui serviront à créer par simulation un bras de contrôle virtuel. Le critère principal de jugement dans ces deux études est l'évolution du **score global de coloration de la cornée et de la conjonctive**, un mois après l'initiation du traitement.

## Objectif

Le but de ce projet est de vérifier si la comparaison indirecte des traitements N et S mène à la même conclusion que l'investigation clinique principale, quant à la non-infériorité de N par rapport à S en se basant sur le même critère d'évaluation principal et en utilisant le même seuil de non-infériorité.

Les deux méthodes appliquées sont l'**appariement optimal** et la **pondération IPTW** (*Inverse Probability of Treatment Weighting*), méthodes qui reposent toutes deux sur le **score de propension**.

## Méthodologie

### Constitution du bras contrôle externe

Le **résumé** de l'étude externe (publiée sur ClinicalTrials.gov, Identifiant NCT01568593) a fourni les données agrégées descriptives : âge, sexe, ancienneté du diagnostic et l'évolution de la coloration de la cornée et de la conjonctive après 1 mois de traitement. Ces données externes agrégées ont été utilisées pour **simuler les données individuelles du bras contrôle externe**, de manière à ce qu'il reflète ces caractéristiques. Cette simulation est réalisée avec le package *Wakefield* du logiciel R. Ainsi, nous obtenons deux bras avec des données individuelles, ce qui permet le calcul du score de propension. Nous sommes alors en mesure d'effectuer une comparaison indirecte entre le bras de traitement N issu de l'investigation clinique et le bras de contrôle S virtuel, d'abord avec la méthode d'appariement optimal, puis avec la pondération IPTW (utilisation des packages R *Opmatch* et *MatchIt*).

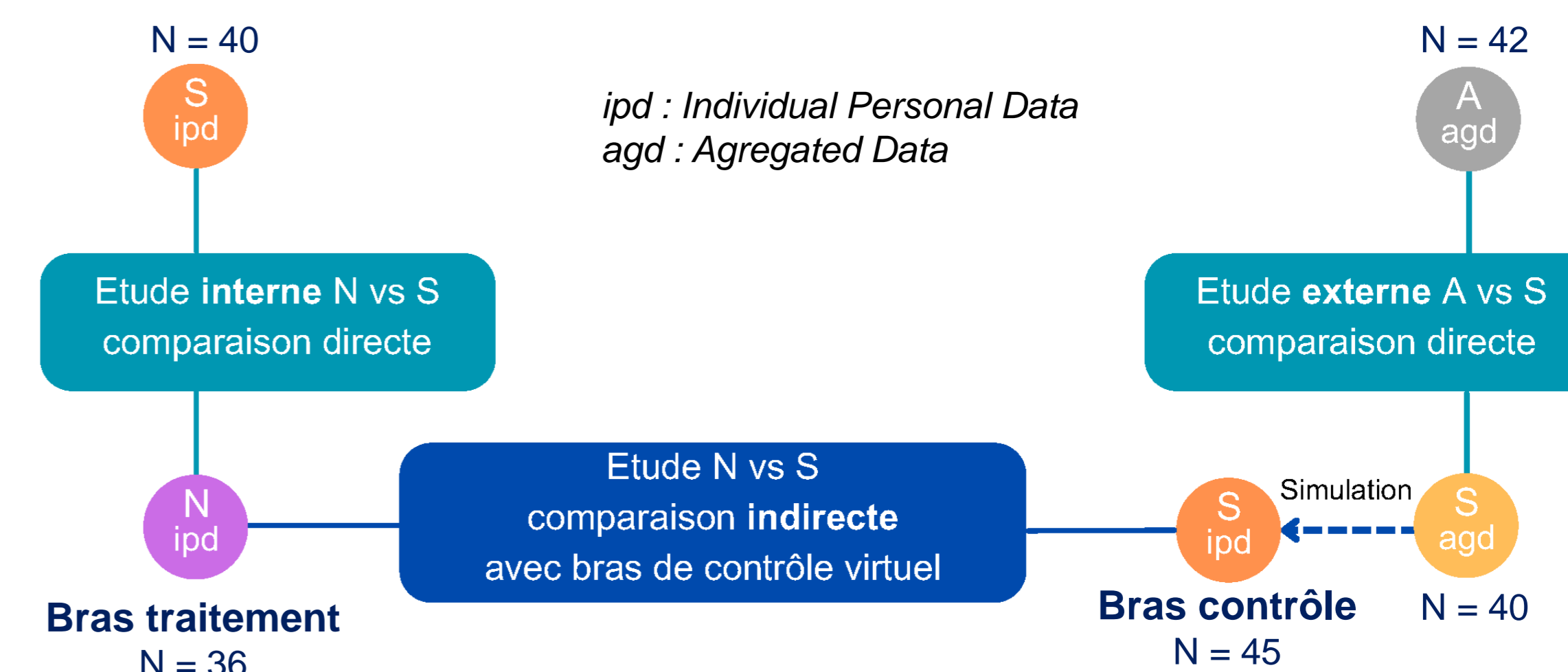


Figure 1. Schéma des relations entre les études

### Calcul du score de propension

Les deux méthodes de comparaison indirecte étudiées ici, l'appariement optimal et la pondération IPTW, s'appuient toutes les deux sur l'estimation du score de propension qui doit être calculé pour chaque individu.

**Score de propension** : probabilité conditionnelle qu'un individu reçoive le traitement d'intérêt (N ici) en fonction d'un ensemble de covariables observées :

$$e(X) = P(T = 1 | X)$$

$e(X)$  : score de propension

$X$  : vecteur de covariables observées

$T = 1$  : l'individu reçoit le traitement d'intérêt ( $T = 0$  sinon)

**Estimation** de ce score par régression logistique dans notre cas :  
 Traitement  $\sim$  Age à l'inclusion + Sexe + Temps depuis le diagnostic

La Figure 2 montre que les médianes du score de propension sont similaires entre les deux groupes, et que ses valeurs sont équilibrées.

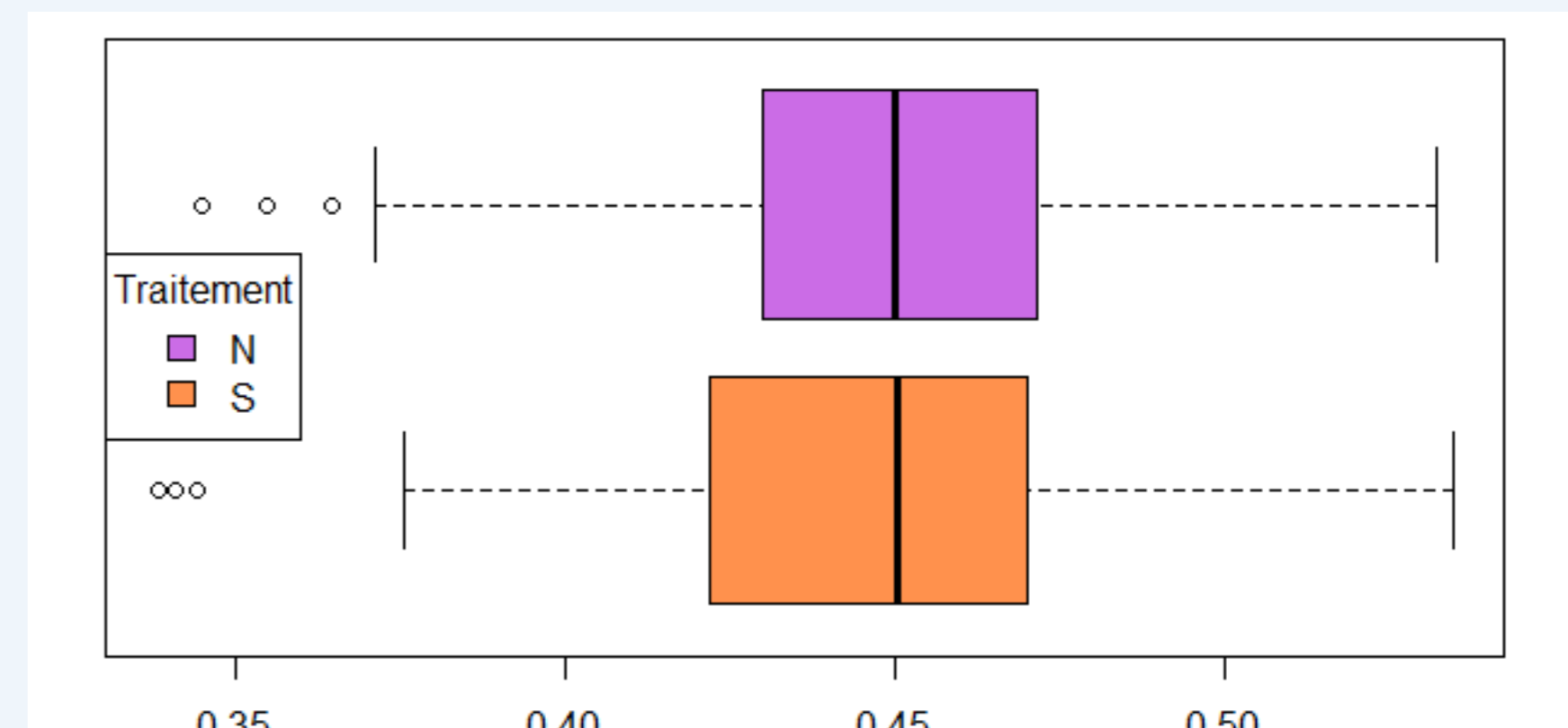


Figure 2. Boîte à moustache du score de propension selon le bras traitement (N) ou contrôle (S)

### Appariement optimal

**Principe** : appairer les individus du bras de contrôle et du bras de traitement ayant des valeurs du score de propension qui sont similaires ou proches, et écarter des analyses les individus non appariés. L'appariement optimal vise à minimiser la différence absolue totale des scores de propension de l'ensemble des paires constituées. Les paires sont déterminées simultanément et sans remplacement.

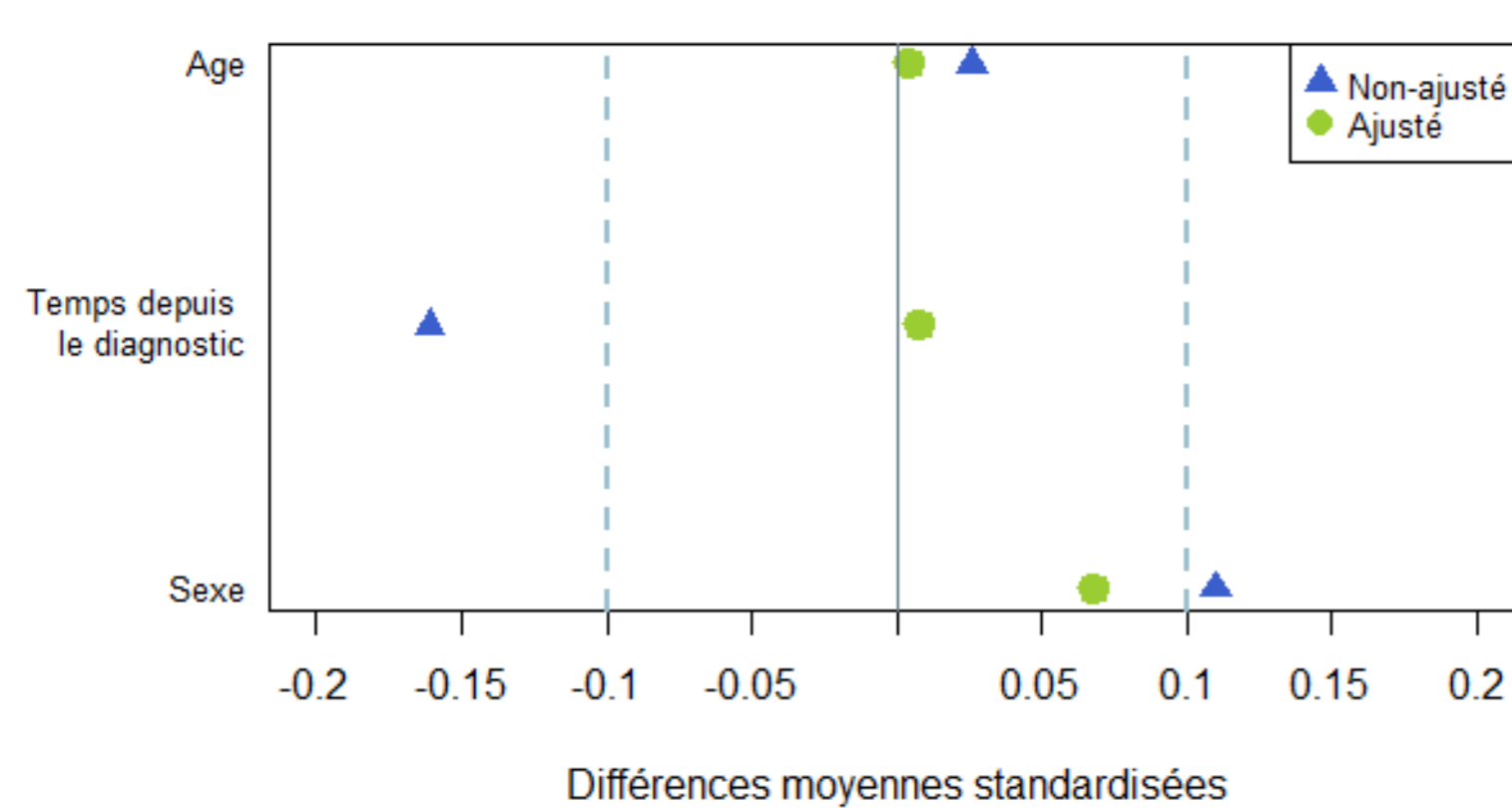


Figure 3. Différences moyennes standardisées, avant et après appariement optimal

Ces deux méthodes permettent un **équilibre des covariables** entre les deux groupes, comme le montrent ces deux figures où les différences de moyennes standardisées après ajustement sont inférieures au **seuil de  $\pm 0.1$**  pour les 3 covariables considérées. Nous pouvons maintenant effectuer les analyses appropriées pour tester la non-infériorité du traitement N par rapport à S.

### Pondération IPTW

**Principe** : attribuer des poids à chaque individu, poids inversement proportionnels à la probabilité d'avoir reçu le traitement d'intérêt (N). Les individus dont cette probabilité est faible se voient attribuer des poids plus élevés que ceux dont cette probabilité est plus forte, qui se voient attribuer des poids plus faibles.

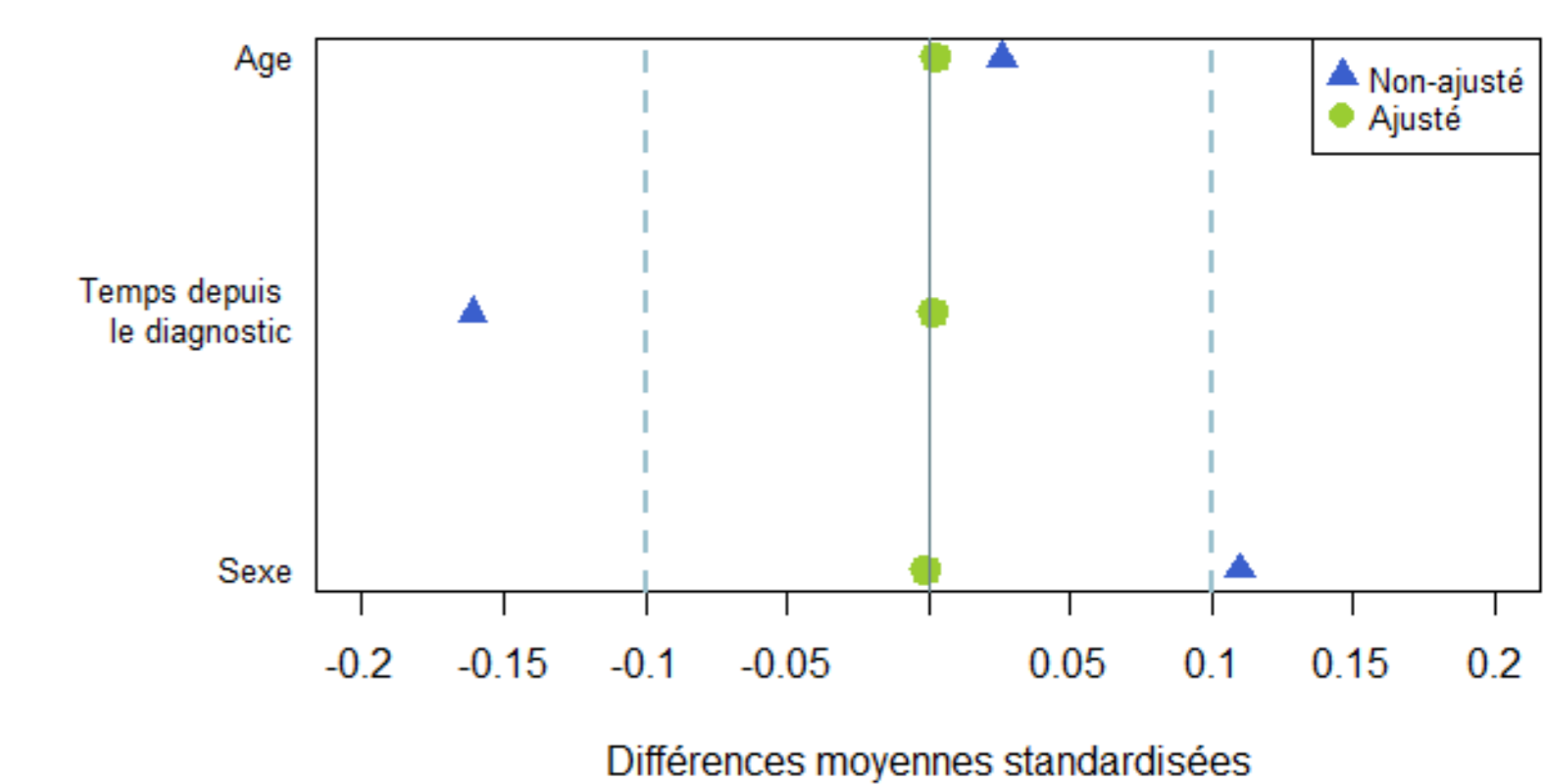


Figure 4. Différences moyennes standardisées, avant et après pondération IPTW

## Résultats

Méthode de comparaison	$\Delta$ [IC95%]
Investigation clinique principale : N vs S	-0.25 [-0.93, 0.43]
Appariement optimal : N vs S (bras virtuel)	0.33 [-0.34, 1.00]
Pondération IPTW : N vs S (bras virtuel)	0.48 [-0.13, 1.09]

$\Delta = \mu_N - \mu_S$ , avec  $\mu_N$  et  $\mu_S$  les différences moyennes entre le score global de coloration à l'inclusion et un mois après le début du traitement dans les bras N et S respectivement.

Avec un seuil de non-infériorité fixé à  $\delta = -2$ , la **non-infériorité** de N est bien démontrée avec les deux méthodes de comparaison indirecte utilisées. En effet, les bornes inférieures des intervalles de confiance à 95% sont supérieures à  $-2$ . Les intervalles de confiances de l'investigation principale ainsi que ceux obtenus par ces méthodes sont représentés sur la Figure 5.

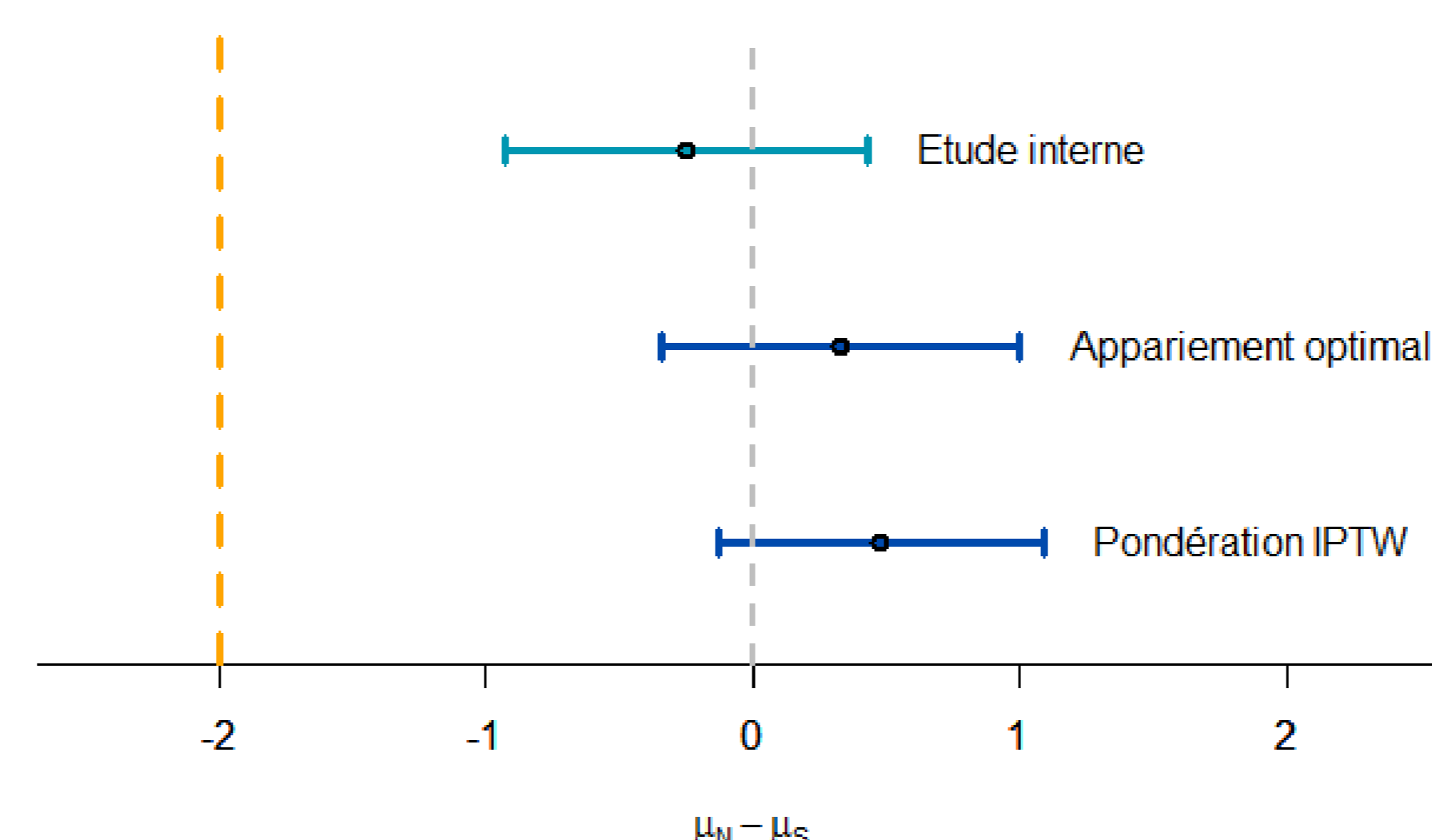


Figure 5. Intervalles de confiance pour le test de la non-infériorité, selon l'analyse effectuée

## Conclusion

Les résultats obtenus après simulation d'un bras de contrôle basé sur des données externes agrégées ont conforté les résultats de l'investigation clinique principale concernant la **non-infériorité du nouveau traitement N** par rapport au traitement standard S, et ce quelle que soit l'une ou l'autre des deux méthodes d'appariement et de pondération appliquées.

Ces méthodologies peuvent être extrapolées à tout type d'étude interventionnelle ou de vie réelle, et sont particulièrement intéressantes dans le cas où il est impossible, compliqué ou non-éthique de réaliser une étude randomisée contrôlée.

Remerciements : Eva Puchalski, ENSAI/ICTA PM pour son aide dans la réalisation de ce poster

## Références

- Cheng, David, Rajeev Ayyagari, and James Signorovitch. "The statistical performance of matching-adjusted indirect comparisons: estimating treatment effects with aggregate external control data." (2020): 1806-1833.
- Serret-Larmande, Arnaud, et al. "A methodological review of population-adjusted indirect comparisons reveals inconsistent reporting and suggests publication bias." *Journal of Clinical Epidemiology* (2023)
- Austin, Peter C. "An introduction to propensity score methods for reducing the effects of confounding in observational studies." *Multivariate behavioral research* 46.3 (2011): 399-424.
- Quantin, Simon. *Estimation avec le score de propension sous R*. Institut National de la Statistique et des Etudes Economiques, 2018.
- Thorlund, Kristian, et al. "Synthetic and external controls in clinical trials—a primer for researchers." *Clinical epidemiology* (2020): 457-467.
- Lambert, Jérôme, et al. "Enriching single-arm clinical trials with external controls: possibilities and pitfalls." *Blood advances* 7.19 (2023): 5680-5690.