

# Essais traditionnels avec des designs randomisés par grappes

MedianaDesigner package

## 1. Introduction

Ce document fournit une description de la méthodologie statistique utilisée dans le module qui prend en charge les essais traditionnels avec des designs randomisés par grappes (fonction `ClustRand`).

Pour plus d'informations sur le paquet `MedianaDesigner`, veuillez-vous référer aux pages internet suivantes :

<http://www.mediana.us/medianadesigner>

<http://medianasoft.github.io/MedianaDesigner>

## 2. Essais randomisés par grappes

### 2.1. Design de l'essai clinique

Les designs randomisés par grappes ont été élaborés pour les cas où les designs d'essai avec randomisation individuelle sont inappropriés ou peu pratiques. Dans ce contexte, des groupes de sujets ou de patients, appelés grappes, sont assignés au hasard à des traitements/interventions expérimentaux ou à des groupes témoins. Les grappes sont généralement définies en termes de sites cliniques dans les essais cliniques ou par différents types de communautés telles que les ménages dans les études d'intervention communautaire. Il est important de noter que les sujets ou les patients au sein de chaque grappe sont considérés comme échangeables. Pour de plus d'informations sur les essais randomisés par grappes, veuillez-vous référer à Hayes et Moulton (2009) et Ahn, Heo et Zhang (2015).

Une évaluation des effets du traitement dans les essais randomisés par grappes est le plus souvent effectuée à l'aide des deux méthodes suivantes :

- Équations d'estimation généralisées (GEE).
- Modèles linéaires généralisés à effets mixtes (GLMEM).

Dans le cadre des GLMEM, les effets aléatoires des grappes et, par conséquent, la variation entre les grappes sont explicitement modélisés. Des méthodes basées sur la vraisemblance

("likelihood") sont ensuite utilisées pour estimer les paramètres d'intérêt tels que l'effet global du traitement dans un essai. Les GEE s'appuient sur une approche alternative où la variation entre les grappes n'est pas explicitement modélisée et où l'attention est centrée sur les corrélations à l'intérieur des grappes.

Les deux méthodes tiennent compte du fait que les résultats pour les patients au sein d'une même grappe sont corrélés. La force de cette corrélation est quantifiée à l'aide du coefficient de corrélation intracluster (ICC). L'ICC est une estimation de la variabilité relative au sein des grappes et entre elles, c'est-à-dire qu'il mesure le niveau auquel les résultats des patients dans une grappe donnée sont similaires les uns aux autres par rapport aux résultats des patients des autres grappes. Un ICC plus faible indique que les résultats des patients au sein de chaque grappe sont faiblement corrélés, ce qui améliore finalement la probabilité de succès de l'essai.

Un design d'essai traditionnel sans analyses intermédiaires est considéré dans ce module. La version actuelle du module prend en charge les essais en groupes parallèles pour évaluer le profil d'efficacité de plusieurs schémas posologiques d'un traitement expérimental ou de plusieurs types d'interventions par rapport à un témoin commun, par exemple un placebo. Le critère d'évaluation principal de l'efficacité dans cet essai peut être continu (normalement distribué) ou binaire et l'analyse de l'efficacité est basée sur une évaluation de la supériorité.

Les considérations générales concernant les designs randomisés par grappes sont présentées dans la section 2.2. Les designs d'essais randomisés par grappes basés sur les méthodes GEE et GLMEM sont discutés dans la section 2.3. Une revue des évaluations de sensibilité prises en charge par ce module est fournie dans la section 2.4. Le processus de design d'essais randomisés par grappes est illustré dans la section 3.

## 2.2. Considérations générales

Pour générer les résultats pour les patients dans les essais randomisés par grappes avec un critère d'évaluation continu, la fonction `ClustRand` repose sur une distribution normale multivariée qui incorpore des effets aléatoires par grappes. Pour les essais avec un critère d'évaluation binaire, une méthode bêta-binomiale est utilisée pour générer des ensembles de grappes avec des résultats binaires corrélés. Cette méthode a été introduite dans Lee et Dubin (1994) et a été utilisée dans de nombreuses études.

Si plusieurs traitements expérimentaux ou interventions sont comparés à un témoin, un ajustement de multiplicité est appliqué aux comparaisons traitement-contrôle afin de préserver le taux d'erreur global de première espèce dans l'essai. La fonction `ClustRand` prend en charge les procédures de test multiples les plus couramment utilisées pour garantir le contrôle du taux d'erreur :

- Procédure de Bonferroni.
- Procédure de Holm.

- Procédure à séquence fixe.
- Procédure de Hochberg.

Pour plus d'informations sur les procédures de tests multiples prises en charge par le paquet `MedianaDesigner`, consultez la documentation de la fonction `MultAdj`.

Une définition standard de la puissance, connue sous le nom de puissance disjonctive, est utilisée dans les essais avec plusieurs comparaisons traitement-contrôle. Dans ce cas, la puissance est définie comme la probabilité qu'au moins un traitement apporte une amélioration par rapport au contrôle.

### 2.3. Design d'essais randomisés par grappes utilisant des GEEs et des GLMEMs

Bien que la méthode GEE soit couramment utilisée dans la conception d'essais randomisés par grappes, il est important de garder à l'esprit que la mise en œuvre standard des GEEs peut conduire à des inférences biaisées. Plus précisément, la mise en œuvre standard repose sur l'estimateur sandwich de la variance et cet estimateur est biaisé lorsque le nombre de grappes est faible (p. ex., moins de 40), ce qui entraîne une inflation du taux d'erreur de première espèce. En général, le degré d'inflation du taux d'erreur de première espèce dans un essai randomisé par grappes ne dépend que du nombre de grappes, par exemple, lorsqu'il y a un petit nombre de grappes, le taux d'erreur de première espèce n'est pas affecté par la taille de la grappe ou l'ICC. Pour atténuer ce problème, la fonction `ClustRand` prend en charge d'autres approches pour définir l'estimateur de la variance pour les GEEs, y compris l'estimateur corrigé du biais proposé par Kauermann et Carroll (2001) et Mancl et DeRouen (2001). Il est à noter que l'estimateur développé par Fay et Graubard (2001) est généralement similaire à la correction de Kauermann-Carroll (Scott et al., 2017) et n'est pas implémenté dans cette fonction.

De plus, la méthode GLMEM est également très populaire pour modéliser les résultats corrélés des patients au sein de chaque grappe. Cette méthode repose sur des modèles à effets mixtes avec un terme aléatoire pour les grappes. Pour effectuer une évaluation fiable des effets du traitement avec la méthode GLMEM, il est important d'utiliser des corrections pour petits échantillons basées sur les degrés de liberté du dénominateur de Kenward-Rogers. Cette approche réduit l'inflation de l'erreur de première espèce dans les essais comportant de petits nombres de grappes et de faibles valeurs d'ICC (Johnson et al., 2015).

### 2.4. Évaluations de sensibilité

Une caractéristique importante de la fonction `ClustRand` est que, en plus des calculs de puissance standard, elle prend en charge un large éventail d'évaluations de sensibilité pour résoudre les problèmes potentiels liés à l'inflation du taux d'erreur de première espèce (voir la section 2.3) et quantifier explicitement l'impact des facteurs clés sur la probabilité de succès :

- ICC mal spécifié.
- Variabilité de la taille des grappes.

- Petit nombre de grappes.

Les évaluations de sensibilité aident le sponsor de l'essai à concevoir des essais plus robustes et à prévenir les problèmes qui surviennent souvent en raison d'une planification inadéquate, tels qu'une puissance réduite en raison de la variabilité de la taille des grappes ou d'un taux d'erreur de première espèce trop élevé.

Après le calcul initial de la puissance, la fonction ClustRand permet à l'utilisateur d'effectuer les évaluations de sensibilité suivantes :

- Une plage de valeurs ICC peut être envisagée.
- Un nombre inégal de grappes peut être pris en compte dans les deux bras de l'essai.
- Des tailles de grappes inégales peuvent être prises en compte dans chaque bras de l'essai.
- Des grappes de taille aléatoire peuvent être envisagées.

Les trois premières évaluations sont assez explicites et la dernière évaluation suppose que la taille de la grappe peut varier d'une simulation à l'autre, ce qui est très courant dans les essais réels randomisés par grappes. Pour effectuer une analyse de sensibilité avec des grappes de taille aléatoire, l'utilisateur spécifie la taille attendue de chaque grappe et les tailles des grappes sont générées à partir d'une distribution de Bernoulli généralisée. Notez que la taille attendue n'a pas besoin d'être constante entre les grappes et que l'utilisateur peut donc évaluer l'impact de tailles inégales des grappes sur les caractéristiques fonctionnelles de l'essai.

### 3. Étude de cas

Un essai clinique randomisé par grappes avec un critère d'évaluation binaire est utilisé pour illustrer les calculs de puissance pris en charge par le module ClustRand. Un design à deux bras avec une dose unique d'un traitement expérimental par rapport à un placebo est envisagé et la méthode GEE est utilisée.

Les hypothèses suivantes ont été formulées dans cette étude de cas. Un ICC commun de 0.3 a été supposée dans les deux bras de l'essai. Les taux de réponse ont été fixés à 30% et 60% dans les groupes placebo et traitement, respectivement (dans ce contexte, un taux de réponse plus élevé indique un profil d'efficacité bénéfique pour ce critère d'évaluation).

Un calcul initial de la puissance repose sur une configuration de base avec un nombre égal de grappes dans les deux bras et des grappes de taille égale. Plus précisément, 16 grappes de 10 patients chacun ont été pris en compte dans chaque bras, ce qui donne la taille de l'échantillon de 160 patients par bras. Les calculs de puissance pour le design randomisé par grappes résultant ont été effectués à l'aide de 10,000 simulations avec un taux d'erreur nominal unilatéral de première espèce défini à 0.025. Les principales caractéristiques fonctionnelles du design randomisé par grappes sont résumées dans le tableau 1. Ce tableau présente le taux d'erreur de

première espèce et la puissance calculée à l'aide des trois estimateurs de variance pris en charge par la fonction ClustRand. L'estimateur sandwich est l'estimateur de variance standard utilisé dans les GEEs et le niveau de puissance associé pour le design sélectionné est de 81%. Cependant, cet estimateur est connu pour sous-estimer la variance réelle dans les petits échantillons et conduit souvent à l'inflation du taux d'erreur de première espèce. En effet, le tableau 1 montre que le taux d'erreur unilatéral de l'estimateur sandwich est bien supérieur à la valeur cible de 0.025. Ce tableau présente également les résultats du calcul du taux d'erreur et de la puissance pour les estimateurs de variance avec correction du biais mis au point par Kauermann et Carroll (2001) et Mancl et DeRouen (2001). Avec ces méthodes, le taux d'erreur unilatéral de première espèce de l'estimateur Mancl-DeRouen est assez proche de la valeur cible de 0.025. Notez que le niveau de puissance associé est inférieur à 80% de quelques points. Si le niveau cible de puissance est fixé à 80 %, le sponsor de l'essai devra augmenter la taille de l'échantillon pour atteindre la probabilité de succès souhaitée dans l'essai.

En plus des caractéristiques fonctionnelles présentées dans le tableau 1, la fonction ClustRand calcule des statistiques descriptives telles que les taux de réponse de l'échantillon et leurs quantiles, par exemple les quartiles inférieur et supérieur, tout au long des simulations. Ces quantités aident l'utilisateur à évaluer dans quelle mesure les valeurs simulées suivent les valeurs supposées des paramètres d'effet du traitement.

Le design de base considérée ci-dessus repose sur un certain nombre d'hypothèses qui peuvent être violées dans l'essai réel, par exemple,

- La taille des grappes est fixée sur l'ensemble des grappes.
- Toutes les grappes ont la même taille.
- Un nombre égal de grappes est supposé dans les deux bras de l'essai.

Comme souligné dans la section 2.4, il est conseillé d'évaluer l'incidence de ces hypothèses sur les conclusions en effectuant des évaluations de sensibilité en ce qui concerne les facteurs clés qui influencent la probabilité de succès dans les essais randomisés par grappes. Il est utile d'examiner l'impact de chaque facteur ou d'une combinaison de facteurs sur les valeurs de puissance pour finalement arriver à un design d'essai robuste.

À titre d'illustration, une analyse de sensibilité avec des tailles de grappes aléatoires a été envisagée. Les tailles de grappes sont échantillonnées à partir d'une distribution de Bernoulli généralisée et, pour fixer les paramètres de la distribution, les tailles de grappes attendues doivent être spécifiées. Par exemple, comme point de départ, on pourrait supposer que les grappes de l'essai ont la même taille moyenne, c'est-à-dire 16 grappes dans chaque bras avec la taille prévue de 10. Le tableau 2 résume les résultats de cette analyse de sensibilité. Il présente la probabilité de succès (puissance) de la méthode GEE avec les estimateurs de Kauermann-Carroll et Mancl-DeRouen. Comme le montre ce tableau, même si des tailles de grappes hétérogènes sont prévues, la puissance de l'essai reste très élevée. En fait, la puissance n'a chuté que d'un

seul point de pourcentage par rapport au tableau 1 où un schéma de regroupement par grappe fixe était supposé. Dans le prolongement de cette évaluation de la sensibilité, l'utilisateur pourrait envisager des cas où la taille attendue varie d'une grappe à l'autre. Des paramètres comme celui-ci imitent des essais réels où il est peu probable que les grappes aient la même taille.

## Références

- Ahn, C., Heo, M., Zhang, S. (2015). Sample Size Calculations for Clustered and Longitudinal Outcomes in Clinical Research. Chapman and Hall/CRC.
- Fay, M.P., Graubard, B.I. (2001). Small sample adjustments for Wald type tests using sandwich estimators. *Biometrics*. 57, 1198-1206.
- Hayes, R.J., Moulton, L.H. (2009). Cluster Randomised Trials: A Practical Approach. Chapman and Hall/CRC.
- Johnson, J.L., Kreidler, S.M., Catellier, D.J., Murray, D.M., Muller, K.E., Glueck, D.H. (2015). Recommendations for choosing an analysis method that controls Type I error for unbalanced cluster sample designs with Gaussian outcomes. *Statistics in Medicine*. 34, 3531-3545.
- Kauermann, G., Carroll, R.J. (2001). A note on the efficiency of sandwich covariance matrix estimation. *Journal of the American Statistical Association*. 96, 1387-1396.
- Lee, E.W., Dubin, N. (1994). Estimation and sample size considerations for clustered binary responses. *Statistics in Medicine*. 13, 1241-1252.
- Mancl, L. A., DeRouen, T.A. (2001). A covariance estimator for GEE with improved small sample properties. *Biometrics*. 57, 126-134.
- Scott, J.M., DeCamp, A., Juraska, M., Fay, M.P., Gilbert, P.B. (2017). Finite-sample corrected generalized estimating equation of population average treatment effects in stepped wedge cluster randomized trials. *Statistical Methods in Medical Research* 26. 583-597.

**Tableau 1. Caractéristiques générales du design randomisé par grappes basée sur la méthode GEE (calculs de puissance initiaux)**

<b>Estimateur de variance</b>	<b>Paramètre</b>	<b>Valeur</b>
Sandwich	Taux d'erreur unilatéral de première espèce	0.033
	Puissance	0.81
Kauermann-Carroll	Taux d'erreur unilatéral de première espèce	0.028
	Puissance	0.792
Mancl-DeRouen	Taux d'erreur unilatéral de première espèce	0.024
	Puissance	0.775

**Table 2. Puissance du design randomisé par grappes basée sur la méthode GEE (évaluation de la sensibilité avec des tailles de grappes aléatoires)**

<b>Estimateur de variance</b>	<b>Paramètre</b>	<b>Valeur</b>
Kauermann-Carroll	Puissance	0.783
Mancl-DeRouen	Puissance	0.762

Il est supposé que les grappes aient la même taille moyenne (16 grappes dans chaque bras avec une taille moyenne de 10).