

Plan d'expérience - Minimization

Rachid Sahli | Abdoul Rafio Bangagne | Antoine Robin - (IUT de Paris- Rives de Seine - Université Paris Cité)

1 | INTRODUCTION

Dans les essais randomisés, la comparabilité des groupes de traitement est assurée par l'attribution des traitements aux patients au moyen d'un mécanisme impliquant un élément aléatoire.

La randomisation est un processus statistique dans lequel un mécanisme aléatoire est utilisé pour sélectionner un échantillon dans une population ou affecter des sujets à différents groupes.

Plusieurs méthodes existent présentant chacune des avantages et des inconvénients. Dans cette étude, nous nous intéresserons à la méthode de minimisation de Pocock et Simon (1975). Nous comparerons les résultats obtenus via cette méthode avec une méthode de randomisation simple. Pour cela, nous utiliserons les tests statistiques de Fisher et de Cochran-Mantel-Haenszel (CMH). Les méthodes de randomisation sont majoritairement utilisées dans le domaine de la santé. Par exemple lors d'essai thérapeutiques ou d'essais clinique avec l'attribution au hasard d'un placebo.

Les principaux objectifs de la randomisation sont d'éviter les biais de sélection, de garantir équilibre et imprédictibilité dans l'attribution de traitement. L'objectif est de permettre la comparabilité initiale des groupes.

2 | METHODE

La randomisation simple permet à chaque participant de l'étude d'avoir une chance égale d'être assigné à l'un des deux groupes. C'est une méthode simple à mettre en œuvre qui consiste à attribuer un traitement à chaque nouveau patient sans tenir compte des facteurs de base dont on sait qu'ils ont une importance pronostique. Cette méthode permet certes une réduction du biais, une application facile et garantit l'imprédictibilité. Mais elle souffre d'une conséquence grave qui est un grave déséquilibre entre les groupes de traitement. Cela est dû à la non-prise en compte de certains facteurs pronostiques, ce qui invalide les résultats de l'essai ou nécessite une analyse secondaire.

La minimisation (Pocock SJ, Simon R, 1975) calcule en temps réel l'attribution du groupe du patient. Cette allocation tend à répartir les données de manière à minimiser les différences entre les groupes en ce qui concerne les facteurs pronostiques. Cette méthode est déterministe, car il est possible de prédire l'attribution du traitement en connaissant préalablement les niveaux des facteurs (variables). Pour éliminer cette possibilité de prédiction, il est nécessaire d'intégrer des éléments d'aléatoires dans les algorithmes de minimisation. En résumé, l'objectif de la minimisation est de se prémunir contre les biais accidentels liés à plusieurs facteurs pronostiques en les prenant en compte simultanément.

L'implémentation complexe, le besoin de disposer de données, et la prévisibilité potentielle font partie des inconvénients de cette méthode. Cependant, elle permet contrairement à la randomisation simple d'avoir un équilibre entre les groupes. Elle est également adaptable aux petits échantillons et peut prendre un grand nombre de facteurs pronostiques connus.

3 | PLAN DE SIMULATION

Pour mener notre étude, nous avons créé deux bases de données. La première comprend 20 patients, tandis que la seconde en compte 50. Elles ont été construites de la même manière.

Nous avons d'abord généré de manière aléatoire les âges de chaque client, en veillant à ce qu'ils suivent une distribution normale (moyenne : 65, écart-type : 10). Ensuite, nous avons attribué un pays à chaque client. La variable "pays" peut prendre l'une des trois valeurs [1, 2, 3], générée également de manière

aléatoire pour chaque patient avec les probabilités suivantes : 1 (50%), 2 (30%), et 3 (20%). De plus, chaque patient dispose d'un identifiant patient.

Nous avons ensuite construit notre algorithme de minimisation sur R en respectant la méthode de Pocock et Simon. Nous avons également réalisé un algorithme pour la méthode de randomisation simple. Par la suite, la simulation de ces deux méthodes nous conduit à créer la variable traitement pour l'attribution du groupe de chaque patient. Le code est disponible sur ce dépôt GitHub : <https://github.com/Rskaki/Minimization>

Une fois ces deux méthodes créées, nous avons pu simuler 1000 études aléatoires pour les deux méthodes sur chaque base de données, avec des ensembles de patients différents à chaque itération. Nous avons donc obtenu 20 000 études pour le fichier contenant 20 patients et 50 000 études pour le fichier contenant 50 patients.

Il est important de préciser que la minimisation étant une méthode déterministe, nous avons alors fait varier un facteur aléatoire à l'attribution du traitement. Ce facteur pouvait être de (60 %, 70 %, 80 % ou 100 %). Par exemple, avec un facteur aléatoire à 80%. Lorsque l'algorithme décide d'attribuer le traitement A au patient suivant, ce dernier n'a en réalité que 80 % de chances de l'obtenir. Les 20% restant sont attribués au traitement B.

Une fois toutes nos simulations réalisées. C'est-à-dire le moment où chaque patient de l'étude dispose d'un traitement. Nous créons une variable statut qui attribue un succès au traitement selon des probabilités suivantes.

Pays du patient	Traitement du succès	
	A	B
1	50 %	30 %
2	30 %	15 %
3	10 %	10 %

Enfin, nous utilisons le test de Fisher, afin de comparer la relation entre l'attribution du traitement et la variable statut, en comparant les p-valeurs obtenues à travers chaque méthode de randomisation. Nous utilisons également le test CMH pour déterminer s'il existe un lien entre le traitement et son statut en contrôlant la variable Pays.

4 | MESURE DES DESEQUILIBRES

Pour chaque méthode, nous avons sommé par étude, le nombre d'individus possédant le traitement A et le traitement B. La différence d'individus entre ces deux groupes constitue le déséquilibre.

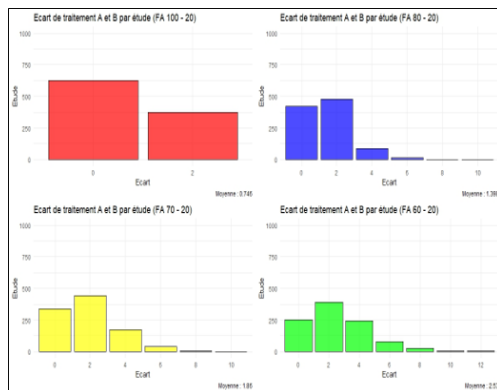


Figure 1 : Ecart Minimisation (taille 20)

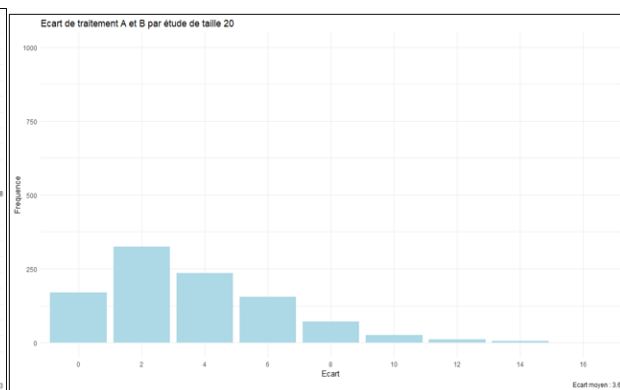


Figure 2 : Ecart Randomisation Simple (taille 20)

Le déséquilibre est un indicateur pertinent ici, car il nous permet de mesurer l'attribution des traitements au sein de chaque groupe.

Pour les échantillons de taille 20, nous observons que la minimisation conduit à des écarts moins importants avec un facteur aléatoire fixé à 100. Environ 65% des études présentent une répartition parfaite, où 10 patients sont assignés à chaque groupe (10A et 10B). Cela indique un équilibre parfait des groupes dans la plupart des cas. Cependant, lorsque le facteur aléatoire diminue à 80%, nous constatons que le nombre d'études parfaitement équilibrées diminue légèrement. Au fur et à mesure de la baisse du facteur aléatoire, les écarts deviennent plus marqués, ce qui indique une moins grande homogénéité dans la répartition des patients entre les groupes. Dans le cas de la randomisation simple, seulement 20% des études présentent une répartition équilibrée, et nous observons des écarts pouvant atteindre jusqu'à 16 patients entre les groupes. Ces écarts importants indiquent une moins bonne répartition des patients entre les traitements. Il est donc nécessaire de passer à des échantillons de taille plus importante pour confirmer ces observations et évaluer si ces tendances se maintiennent. En raison de la petite taille de l'échantillon, il est difficile de dégager une tendance significative.

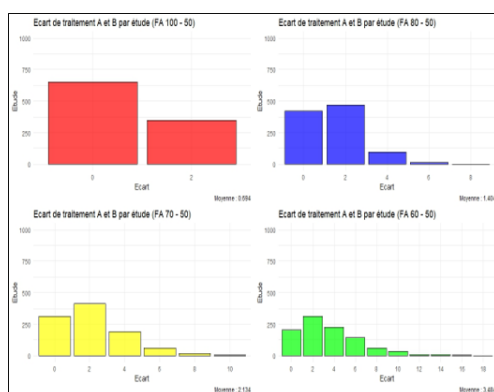


Figure 3 : Ecart Minimisation (taille 50)

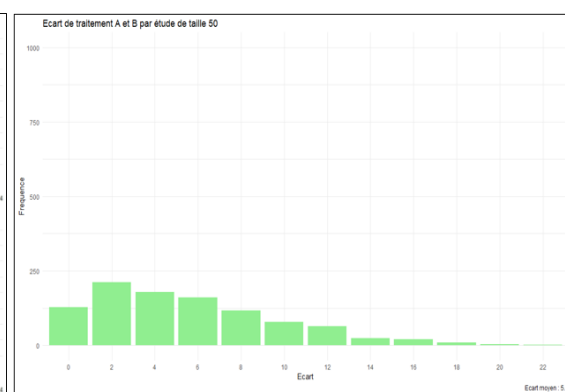


Figure 4 : Ecart Randomisation Simple (taille 50)

Pour les échantillons de taille 50, la tendance observée pour la randomisation simple et la minimisation reste la même. Dans le cas de la minimisation, quelle que soit la valeur du facteur aléatoire, la moyenne des écarts entre les groupes de traitement reste inférieure à celle observée avec la randomisation simple. De plus, aucune étude ne présente un écart supérieur à 18 pour la minimisation, tandis que pour la randomisation simple, des écarts allant jusqu'à 22 sont observés. Ces observations soulignent la capacité

de la minimisation à maintenir un meilleur équilibre entre les groupes de traitement, même lorsque le facteur aléatoire diminue.

Au regard de tout ce qui précède concernant le déséquilibre, il est clair que plus le facteur aléatoire diminue, plus les écarts entre les groupes de traitement augmentent pour la minimisation. Cependant, dans le cas où le facteur aléatoire ne serait pas faible, la minimisation parvient à maintenir un meilleur équilibre, contrairement à la randomisation simple où des écarts significatifs sont observés.

Dans notre cas, nous observons que les écarts augmentent à partir du facteur aléatoire de 60%. L'écart moyen de la minimisation avec le facteur aléatoire de 60% se rapproche grandement de celui de la randomisation simple.

Dans le cadre de notre étude, nous avons également analysé les p-valeurs de Fisher et du test CMH obtenu à l'aide des deux méthodes de randomisation.

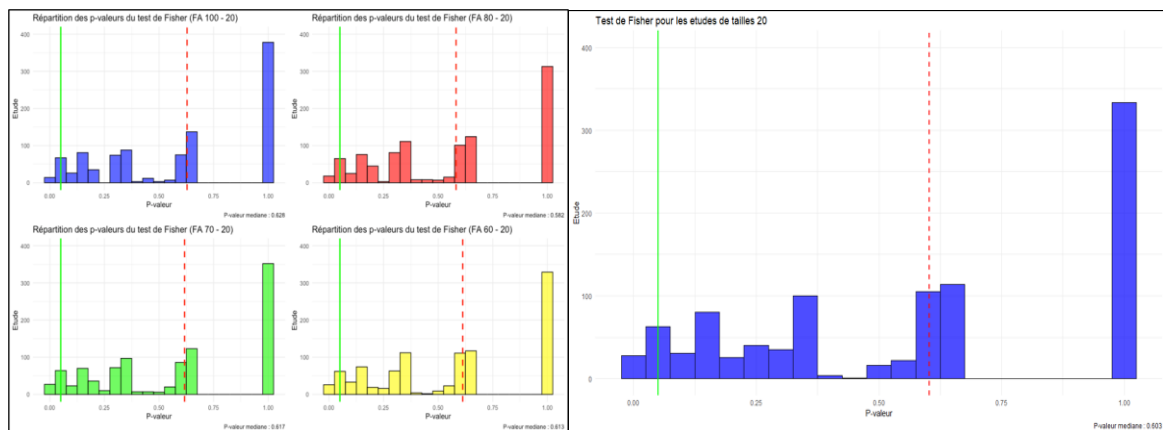


Figure 5 : P-valeur Fisher Minimisation

Figure 6 : P-valeur Fisher Randomisation Simple

Les résultats ne sont pas comparables, car la répartition des p-valeurs est similaire pour les deux méthodes, indépendamment de la taille de l'échantillon (20 ou 50) et des différents facteurs aléatoires. Les p-valeurs médiane des deux méthodes sont également très proche. Cette ressemblance significative est due à la faible taille de l'étude. Avec un nombre plus élevé de patients par étude, la comparaison aurait été plus intéressante. Les résultats obtenus à l'aide du test CMH confirment la même conclusion.

Voici un exemple de la première étude obtenu à l'aide de la minimisation au facteur 100.

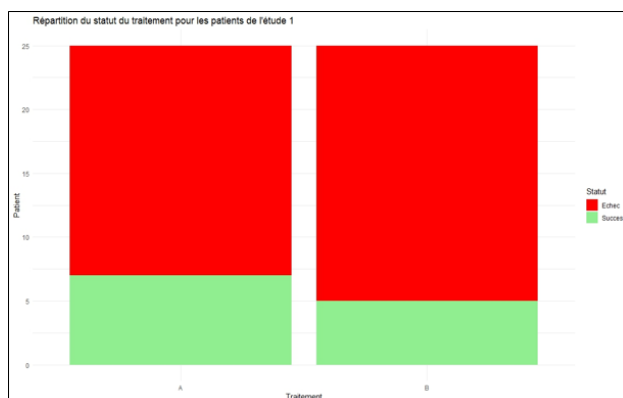


Figure 7 : Répartition du statut

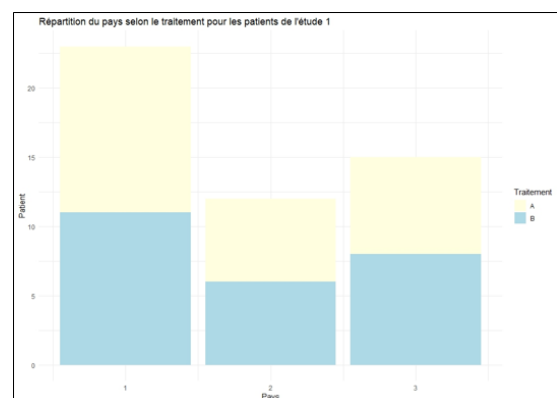


Figure 8 : Répartition du pays

```
> table(etude1$Traitement)
```

A	B
25	25

Figure 5 : Répartition du traitement

Ces résultats confirment bien la réussite de notre algorithme, car ils sont concordants avec les probabilités affectées au début. De plus, on constate que le traitement est réparti équitablement entre les groupes. Ici, 25 patients ont le traitement A et 25 patients ont le traitement B.

5 | RESULTATS

Proportion d'études qui confirme l'indépendance du succès et du traitement :

% p valeur > 0.05	Minimisation (FA 100)	Minimisation (FA 80)	Minimisation (FA 70)	Minimisation (FA 60)	Randomisation simple
Taille 20	95.2	94.8	94.1	95.1	93.49
Taille 50	88.3	85.4	87.8	85.3	81.6

Figure 7 : Comparaison Fisher

% p valeur > 0.05	Minimisation (FA 100)	Minimisation (FA 80)	Minimisation (FA 70)	Minimisation (FA 60)	Randomisation simple
Taille 20	91.4	80.7	75.8	71.5	90.67
Taille 50	67.7	64.5	67.9	68.6	60

Figure 8 : Comparaison CMH

Dans le cas des tests de Fisher, le taux d'acceptation de la proportion d'études confirmant l'indépendance du succès et du traitement est toujours supérieur pour la minimisation par rapport à la randomisation simple, quelle que soit le facteur aléatoire.

Dans le cas des tests CMH, les résultats ne sont pas vraiment pertinents, car lors du traitement, nous avons rencontré de nombreuses p-valeurs manquantes. Les études étant trop petites, les conditions du test CMH n'étaient pas réunies. On constate quand même que le taux d'acceptation de la proportion d'études confirmant l'indépendance du succès et du traitement est toujours inférieur pour la randomisation simple au taux de la minimisation pour les étude de taille 50.

En ce qui concerne les écarts entre les groupes, nous avons constaté dans la partie précédente que la randomisation simple présentait un nombre élevé d'écarts par groupe comparé à la minimisation. L'écart moyen entre les groupes obtenus à l'aide de la minimisation est toujours inférieur à ceux obtenus par la randomisation simple, quelle que soit la taille de l'étude et le niveau du facteur aléatoire de la minimisation. Mais si le niveau du facteur aléatoire diminue, l'écart moyen entre les groupes de la minimisation tend à augmenter.

6 | DISCUSSION / CONCLUSION

A travers cette étude où nous avons comparé la minimisation à la randomisation simple, les résultats obtenus confirment que la minimisation est meilleure.

Les écarts entre les groupes sont beaucoup plus faibles, comme en témoignent les mesures de déséquilibre comparé à ceux de la randomisation simple. La similitude dans la répartition des p-

valeurs des différents tests s'explique par la petitesse des échantillons. Le facteur aléatoire est important, car la minimisation est une méthode déterministe.

La minimisation est une méthode très performante pour de petites études. Reprendre ces études avec un plus grand nombre de patients pourrait être pertinent.

[An Overview of Randomization and Minimization Programs for Randomized Clinical Trials - Prof. Mahmoud Saghaei, PO Box 931, Al-Zahra Medical Center, Isfahan, Iran.](#)

[https://www.ajodo.org/article/S0889-5406\(11\)00790-6/pdf](https://www.ajodo.org/article/S0889-5406(11)00790-6/pdf)