

# Sélection optimale d'une règle d'arrêt pour futilité

MedianaDesigner package

## 1. Introduction

Ce document fournit une description de la méthodologie statistique utilisée dans le module de futilité (fonction FutRule).

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

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

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

## 2. Sélection optimale d'une règle d'arrêt pour futilité

### 2.1. Design de l'essai clinique

Considérons un essai de Phase II ou de phase III à plusieurs bras. L'essai sera mené pour étudier l'efficacité et l'innocuité de plusieurs doses ou schémas posologiques d'un traitement expérimental par rapport à un placebo. Le critère principal d'efficacité dans cet essai clinique peut être basé sur un critère continu, binaire ou de type temps jusqu'à un évènement.

Une seule analyse intermédiaire sera effectuée en levant l'aveugle dans le cadre de cet essai. Une règle d'arrêt pour futilité sera appliquée lors de l'analyse intermédiaire pour effectuer une évaluation de la futilité pour chaque groupe de traitement par rapport au placebo. Un bras de traitement sera arrêté pour futilité si le traitement correspondant est peu susceptible d'être efficace, par exemple, la probabilité prédite de succès à l'analyse finale est faible. Une règle d'arrêt pour futilité est définie dans la Section 2.2 et une règle d'arrêt pour futilité optimale maximisant les taux de sensibilité et de spécificité de cette règle est dérivée dans la Section 2.3. L'approche proposée pour définir des règles optimales d'arrêt pour futilité est illustrée dans la Section 3.

### 2.2. Règle d'arrêt pour futilité

La règle d'arrêt pour futilité lors de l'analyse intermédiaire peut être mis en place à l'aide de n'importe quelle définition pertinente de la probabilité de succès prédite, mais, comme dans les modules de design adaptatif (fonctions ADSSMod, ADTreatSel et ADPopSel), elle sera définie à l'aide de la puissance conditionnelle. Comme dans ces modules, la puissance conditionnelle est

définie comme la probabilité d'établir un effet de traitement significatif à l'analyse finale conditionnelle aux données intermédiaires, veuillez-vous référer à Wassmer et Brannath (2016, Chapitre 7). La dérivation de la puissance conditionnelle est fournie dans la documentation de la fonction ADSSMod et sera donc omise ici. En outre, comme indiqué dans les modules de design adaptatif, les règles d'arrêt pour futilité sont non-contraignantes ("non-binding") et peuvent être ignorées par le sponsor de l'essai clinique ou le comité de surveillance des données.

Notons  $m$  le nombre total de doses ou de schémas posologiques dans l'essai (les doses ou les schémas posologiques seront appelés traitements). En se concentrant sur la comparaison du  $k$ ème traitement par rapport au placebo,  $k = 1, \dots, m$ , la puissance conditionnelle pour cette comparaison traitement-placebo est notée  $CP_k$ .

Un bras de traitement sera abandonné lors de l'analyse intermédiaire de la futilité si la puissance conditionnelle correspondante ne dépasse pas un seuil prédéfini désigné par  $c$ , où  $0 < c < 1$ . Par exemple, le  $k$ ème traitement sera abandonné si

$$CP_k \leq c.$$

L'essai sera interrompu pour futilité si tous les traitements sont abandonnés.

### 2.3. Règle optimale d'arrêt pour futilité

Il est courant de fixer le seuil de futilité ( $c$ ) à une valeur faible telle que 0.2 ou 0.3. Au lieu d'utiliser un seuil arbitraire de ce type, une valeur optimale de  $c$  est facile à dériver en utilisant l'approche standard pour définir les règles de décisions lors d'analyses intermédiaires, voir, par exemple, Chuang-Stein et al. (2011) et Wang et al. (2014). Cette approche repose sur le calcul des taux de sensibilité et de spécificité de la règle d'arrêt pour futilité.

Les taux de sensibilité et de spécificité sont définis comme suit :

- Le taux de sensibilité est défini comme la probabilité de retenir correctement au moins un traitement efficace lors de l'analyse intermédiaire, c'est-à-dire la probabilité que la puissance conditionnelle soit supérieure au seuil de futilité ( $c$ ) dans un ou plusieurs bras de traitement. Cette probabilité est évaluée sous l'hypothèse alternative d'un effet bénéfique, c'est-à-dire que tous les traitements  $m$  sont supposés efficaces.
- Le taux de spécificité est défini comme la probabilité d'abandonner correctement tous les traitements inefficaces lors de l'analyse intermédiaire, c'est-à-dire la probabilité que la puissance conditionnelle soit inférieure au seuil de futilité ( $c$ ) dans les bras de traitement. Cette probabilité est évaluée sous l'hypothèse nulle de l'absence d'effet, c'est-à-dire que tous les traitements  $m$  sont supposés inefficaces.

Le taux de sensibilité peut être considéré comme le taux vrai-positif et le taux de spécificité est égal à un moins le taux faux-positif. Lorsque deux doses ou schémas posologiques (traitements)

ou plus sont évalués dans l'essai, les définitions des taux de vrais-positifs et de faux-positifs tiennent compte de la multiplicité des résultats possibles lors de cette analyse intermédiaire.

On s'attend à ce que les règles de décision démontrent des taux élevés de sensibilité et de spécificité. Afin de construire une règle d'arrêt de futilité optimale, le seuil de futilité est choisi pour maximiser la moyenne des taux de sensibilité et de spécificité. La moyenne des taux de sensibilité et de spécificité est connue sous le nom de taux de précision.

### 3. Étude de cas

Pour illustrer le processus de dérivation d'un seuil optimal pour une règle d'arrêt pour futilité, considérons un essai de phase III chez des patients atteints de schizophrénie. Le profil d'efficacité de deux doses d'un traitement expérimental sera évalué par rapport à un placebo. L'analyse d'efficacité principale sera effectuée en utilisant la variation du score total PANSS (Positive and Negative Syndrome Scale) à la Semaine 6 par rapport à la Baseline.

Un design équilibré avec 160 patients recrutés par groupe sera utilisé dans l'essai et une analyse intermédiaire sera effectuée en levant l'aveugle après que 50% des patients aient terminé la période de traitement de 6 semaines ou abandonné l'essai avant la fin de la période de traitement. Un taux d'abandon de 25% des patients est supposé dans l'essai.

Pour définir une règle optimale d'arrêt pour futilité dans cet essai, le taux de sensibilité a été calculé en supposant que la taille réelle de l'effet dans chaque bras de traitement correspond à la différence cliniquement importante minimale, c'est-à-dire 0.25, et le taux de spécificité a été calculé en supposant que la taille réelle de l'effet dans chaque bras de traitement est de 0.

La Figure 1 montre les taux de sensibilité et de spécificité résultants de l'analyse intermédiaire en fonction du seuil de futilité ( $c$ ). Par définition, le taux de sensibilité est une fonction monotone décroissante du seuil tandis que le taux de spécificité est une fonction monotone croissante du seuil. Une valeur optimale de  $c$  augmente simultanément les taux de sensibilité et de spécificité. Pour trouver cette valeur optimale, le taux de précision doit être calculé.

Le taux de précision est illustré dans la Figure 2 et il est facile de vérifier que le taux de précision est maximisé si le seuil de futilité est fixé à  $c = 0.27$ , c'est-à-dire qu'un traitement ne sera retenu lors de l'analyse intermédiaire que si la puissance conditionnelle est supérieure à 27%, sinon il sera abandonné. Avec ce choix du seuil de futilité, les taux de sensibilité et de spécificité dépassent tous deux 80%. Il résulte également de cette figure que la courbe de précision est assez plate autour de la valeur optimale et que des seuils raisonnablement proches de 0.27 se traduisent par des taux de précision élevés. Par exemple, la figure présente un intervalle optimal de 95% pour  $c$ , qui s'étend de 0.05 à 0.77 et inclut les seuils pour lesquels le taux de précision n'est pas inférieur de plus de 5% au taux de précision optimal (sur une échelle absolue).

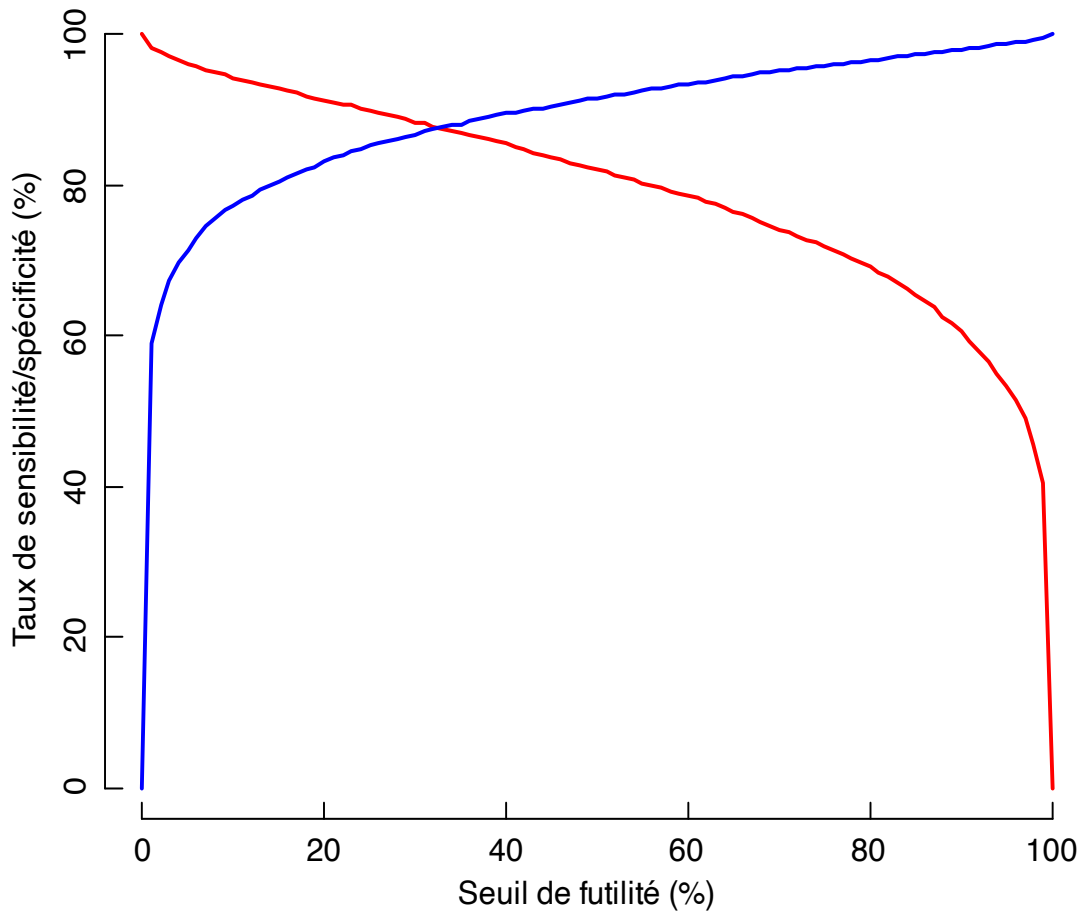
## Références

Chuang-Stein, C., Kirby, S., French, J., Kowalski, K., Marshall, S., Smith, M. K. (2011). A quantitative approach for making go/no-go decisions in drug development. *Drug Information Journal*. 45, 187-202.

Wang, D., Cui, L., Zhang, L., Yang, B. (2014). An ROC approach to evaluate interim go/no-go decision-making quality with application to futility stopping in the clinical trial designs. *New Developments in Statistical Modeling, Inference and Application*. Jin, Z., Liu, M., Luo, X. (editors). Springer, New York. 121-147.

Wassmer, G., Brannath, W. (2016). *Group Sequential and Confirmatory Adaptive Designs in Clinical Trials*. New York: Springer.

**Figure 1. Taux de sensibilité et spécificité en fonction du seuil de futilité**



Courbe rouge : Taux de sensibilité (probabilité de retenir correctement au moins un bras de traitement lors de l'analyse intermédiaire, évaluée sous l'hypothèse alternative d'un effet bénéfique, c'est-à-dire que tous les traitements sont efficaces). Courbe bleue : Taux de spécificité (probabilité d'arrêter correctement tous les bras de traitement lors de l'analyse intermédiaire en raison de la futilité, évaluée sous l'hypothèse nulle de l'absence d'effet, c'est-à-dire que tous les traitements sont inefficaces).

Figure 2. Taux de précision en fonction du seuil de futilité

