

LA MÉTHODE DE SIMON

Spécialisé dans le traitement statistique des études cliniques,
le Département Biométrie de FOVEA vous propose au travers
de ce fascicule quelques notions simples concernant
LA METHODE DE SIMON.

*en
pratique*

Traitement statistique
DES ETUDES CLINIQUES



INTRODUCTION

En recherche clinique, les études de phase II à un seul bras sont utilisées pour détecter de façon précoce, avec un minimum de sujets, des molécules à l'efficacité suffisante pour justifier des études de phase III. L'objectif est de minimiser le risque d'éliminer une molécule active ou a contrario, de rejeter rapidement des molécules insuffisamment actives et d'éviter

ainsi d'y exposer inutilement des patients. Différentes méthodes ont été proposées pour déterminer le nombre de sujets à inclure dans ces études de phase II à un seul bras : Gehan, Fleming, Simon, Ensign. Les méthodes de Gehan et de Fleming ont chacune fait l'objet d'un précédent fascicule. Nous détaillons aujourd'hui la méthode de Simon.

1

Principes généraux

Dans le cas de l'évaluation d'un nouveau médicament avec un seul groupe de sujets traités, l'un des plans statistiques les plus utilisés pour des essais de phase II a été la méthode de Gehan publiée en 1961, conçue pour rejeter précocement des molécules si aucune réponse n'était observée dans une première cohorte de patients. Vingt ans après, des plans plus efficaces ont été développés : Fleming en 1982, Simon en 1989, Ensign en 1995.

Les plans à 2 étapes de Gehan permettent d'interrompre l'inclusion des sujets si aucun succès n'est observé sur les n_1 premiers sujets. Ils sont faciles à utiliser, d'où leur succès, mais présentent l'inconvénient d'exposer beaucoup de sujets à un produit potentiellement inefficace lors de la deuxième étape.

Le plan à plusieurs étapes de Fleming permet un arrêt précoce de l'étape dans le cas où les résultats intermédiaires sont extrêmes, soit en faveur de l'efficacité, soit en faveur de l'inefficacité. Les plans d'Ensign sont des plans à 3 étapes qui combinent la 1^{re} étape de Gehan avec les 2 étapes de Simon.

La méthode de Simon est une version optimisée de la méthode de Gehan. Elle comporte **2 étapes** et, comme la méthode de Fleming, permet **un arrêt précoce dès la première étape** si des résultats indéniables en faveur de l'efficacité ou de l'inefficacité du produit étudié sont observés.

Il existe deux variantes des plans de Simon : **le design Optimal qui minimise le nombre moyen de sujets exposés, le design Minimax**

qui **minimise le nombre maximum de sujets exposés**. Ces designs sont élaborés au moyen d'un **logiciel** et les paramètres sont disponibles dans **des tables**. A la différence des plans de Fleming, le nombre total de sujets et le nombre de sujets à chaque étape ne sont pas définis par l'investigateur, mais résultent des contraintes de **la minimisation**.

2

Méthodologie

Sont définies au préalable les valeurs de la borne p_0 , borne supérieure de la zone d'inefficacité et de la borne p_1 , borne inférieure de la zone d'efficacité.



La réponse au traitement est bimodale, succès ou échec, et suit une loi binomiale. La probabilité réelle de succès est définie par le paramètre p , pourcentage de succès observé. Sont définies également les valeurs des risques d'erreur de 1^{re} espèce α et de 2^e espèce β . En effet, il est aussi important de ne pas rejeter à tort un médicament inefficace (risque α) que de rejeter à tort un médicament efficace (risque β).

Les 2 hypothèses à considérer sont :

$H_0 : p \leq p_0$ (H_0 : hypothèse nulle d'inefficacité)

$H_1 : p \geq p_1$ (H_1 : hypothèse alternative d'efficacité)

Ainsi :

- . Si $p \leq p_0$, la molécule étudiée sera considérée comme insuffisamment efficace et l'étude sera interrompue
- . Si $p \geq p_1$, la molécule étudiée sera considérée comme suffisamment efficace pour entreprendre des études de phase III

De ces 4 paramètres (p_0, p_1, α, β) vont découler **le nombre de sujets à inclure (n_1, n_2) à chaque étape et la valeur seuil du nombre de réponses objectives (r_1, r)** qui vont déterminer la poursuite ou l'arrêt de l'étude après la 1^{re} étape (r_1) et la poursuite du développement en phase III selon la valeur de r . Ces valeurs sont calculées selon les principes de la loi binomiale par un logiciel spécifique qui, pour les paramètres α, β, p_0 et p_1 , élabore **le design Optimal et le design Minimax**.

Le design Optimal sera

construit afin que **la taille moyenne de l'échantillon (EN) soit la plus faible possible** lorsque le taux d'efficacité du traitement étudié est supérieur au taux minimal nécessaire pour poursuivre le développement du produit.

Le design Minimax sera construit pour que **le nombre maximum de sujets exposés soit minimisé**. Si lors de la 1^{re} étape, le nombre de réponses observées est inférieur ou égal à r_1 ou supérieur à r , l'étude est stoppée. Le projet de phase III est abandonné dans le 1^{er} cas, dans le 2^e cas le produit est déclaré admissible pour les études de phase III, sans passer par la 2^e étape.

Si ce nombre de réponses observées est supérieur à r_1 et inférieur ou égal à r , l'étude est poursuivie dans la 2^e étape. A la fin de l'étude, soit le nombre de succès observés est inférieur ou égal à r et le produit est considéré définitivement comme vraisemblablement inefficace, soit il est supérieur à r et le produit est déclaré potentiellement efficace.

3

Interprétation des tables 1 et 2

■ Description

Les tables 1 et 2 présentent les designs correspondant aux valeurs des différents paramètres.

La table 1 est basée sur l'hypothèse $p_1 - p_0 = 0,20$

La table 2 est basée sur l'hypothèse $p_1 - p_0 = 0,15$

Les valeurs du design Optimal sont présentées dans la partie gauche de la table et celles du design Minimax dans la partie droite. Pour chaque couple $p_0 - p_1$, 3 lignes sont disponibles correspondant aux valeurs de α et β , égales respectivement à 0,10 et 0,10, à 0,05 et 0,20, à 0,05 et 0,10.

Les données d'une ligne considérée correspondent à la taille optimale n_1 de l'échantillon pour la 1^{re} étape, la taille maximum n de l'échantillon total (1^{re} + 2^e étape), les limites supérieures du taux de réponse observé qui provoque le rejet du produit étudié, à la fin de la 1^{re} étape (r_1/n_1) ou à la fin de l'étude (r/n). Figurent également sur cette ligne, dans l'hypothèse d'une efficacité $p \leq p_0$, la probabilité $PET(p_0)$ d'arrêt précoce de l'étude et le nombre moyen de sujets à inclure $EN(p_0)$, calculé selon la formule $n_1 + [1 - PET(p_0)] \times n_2$.

■ Exemples (Table 1)

1) Design Optimal

- Pour $\alpha = 0,05, \beta = 0,10, p_0 = 0,10$ et $p_1 = 0,30$ (ligne 6 de la Table 1), dans le cas du design Optimal le plan est le suivant :
- A la 1^{re} étape, il convient d'inclure $n_1=18$ sujets
 - Si l'on observe moins de 3 succès ($r_1 \leq 2$), l'étude est stoppée et on conclut à l'inefficacité du produit
 - Si l'on observe 7 succès et plus ($r_1 > 6$), l'étude est stop-

LA MÉTHODE DE SIMON EN PRATIQUE

pée et on conclut que le produit est admissible pour la phase III sans passer par la 2^e étape

- Si $3 \leq r_1 \leq 6$, 17 sujets supplémentaires sont inclus, soit un total n de 35 sujets, dans l'étape 2

A la fin de cette étape, on conclut à l'inefficacité si l'on observe moins de 7 succès, le produit est rejeté définitivement. A l'inverse, si l'on observe au moins 7 succès, le produit est considéré comme admissible pour un développe-

ment ultérieur en phase III.

La probabilité d'arrêt précoce $PET(p_0)$, dans l'hypothèse d'une efficacité de 10% ($p = 0,10$), est de 0,71 et le nombre moyen de sujets à inclure $EN(p_0) = 22,5$.

2) Design Minimax

Si maintenant nous nous intéressons au design Minimax (ligne 6 de la Table 1), le nombre maximal n de sujets est égal à 33 au lieu de 35 pour le design Optimal (soit -5,7%). En

revanche, le nombre moyen $EN(p_0)$ de sujets est égal à 26,2 au lieu de 22,5 (soit +16,4%).

Ainsi, si l'étude est arrêtée à la 1^{re} étape (moins de 3 succès ou au moins 7 succès), le nombre n_1 de patients exposés est égal à 22 pour le design Minimax au lieu de 18 pour le design Optimal. Néanmoins, si l'on observe entre 3 et 6 succès, on passe à l'étape 2 en incluant 11 sujets supplémentaires, soit 33 sujets au total, versus 35 sujets au total dans le design Optimal.

Table 1 Designs for $p_1 - p_0 = 0.20^a$

| α | β | p_0 | p_1 | Optimal Design | | | | Minimax Design | | | |
|----------|---------|-------|-------|------------------|--------------|-----------|------------|------------------|--------------|-----------|------------|
| | | | | $\leq r_1 / n_1$ | $\leq r / n$ | $EN(p_0)$ | $PET(p_0)$ | $\leq r_1 / n_1$ | $\leq r / n$ | $EN(p_0)$ | $PET(p_0)$ |
| 0.10 | 0.10 | 0.05 | 0.25 | 0/9 | 2/24 | 14.5 | 0.63 | 0/13 | 2/20 | 16.4 | 0.51 |
| 0.05 | 0.20 | | | 0/9 | 2/17 | 12.0 | 0.63 | 0/12 | 2/16 | 13.8 | 0.54 |
| 0.05 | 0.10 | | | 0/9 | 3/30 | 16.8 | 0.63 | 0/15 | 3/25 | 20.4 | 0.46 |
| 0.10 | 0.10 | 0.10 | 0.30 | 1/12 | 5/35 | 19.8 | 0.65 | 1/16 | 4/25 | 20.4 | 0.51 |
| 0.05 | 0.20 | | | 1/10 | 5/29 | 15.0 | 0.74 | 1/15 | 5/25 | 19.5 | 0.55 |
| 0.05 | 0.10 | | | 2/18 | 6/35 | 22.5 | 0.71 | 2/22 | 6/33 | 26.2 | 0.62 |
| 0.10 | 0.10 | 0.20 | 0.40 | 3/17 | 10/37 | 26.0 | 0.55 | 3/19 | 10/36 | 28.3 | 0.46 |
| 0.05 | 0.20 | | | 3/13 | 12/43 | 20.6 | 0.75 | 4/18 | 10/33 | 22.3 | 0.50 |
| 0.05 | 0.10 | | | 4/19 | 15/54 | 30.4 | 0.67 | 5/24 | 13/45 | 31.2 | 0.66 |
| 0.10 | 0.10 | 0.30 | 0.50 | 7/22 | 17/46 | 29.9 | 0.67 | 7/28 | 15/39 | 35.0 | 0.36 |
| 0.05 | 0.20 | | | 5/15 | 18/46 | 23.6 | 0.72 | 6/19 | 16/39 | 25.7 | 0.48 |
| 0.05 | 0.10 | | | 8/24 | 24/63 | 34.7 | 0.73 | 7/24 | 21/53 | 36.6 | 0.56 |
| 0.10 | 0.10 | 0.40 | 0.60 | 7/18 | 22/46 | 30.2 | 0.56 | 11/28 | 20/41 | 33.8 | 0.55 |
| 0.05 | 0.20 | | | 7/16 | 23/46 | 24.5 | 0.72 | 17/34 | 20/39 | 34.4 | 0.91 |
| 0.05 | 0.10 | | | 11/25 | 32/66 | 36.0 | 0.73 | 12/29 | 27/54 | 38.1 | 0.64 |
| 0.10 | 0.10 | 0.50 | 0.70 | 11/21 | 26/45 | 29.0 | 0.67 | 11/23 | 23/39 | 31.0 | 0.50 |
| 0.05 | 0.20 | | | 8/15 | 26/43 | 23.5 | 0.70 | 12/23 | 23/37 | 27.7 | 0.66 |
| 0.05 | 0.10 | | | 13/24 | 36/61 | 34.0 | 0.73 | 14/27 | 32/53 | 36.1 | 0.65 |
| 0.10 | 0.10 | 0.60 | 0.80 | 6/11 | 26/38 | 25.4 | 0.47 | 18/27 | 24/35 | 28.5 | 0.82 |
| 0.05 | 0.20 | | | 7/11 | 30/43 | 20.5 | 0.70 | 8/13 | 25/35 | 20.8 | 0.65 |
| 0.05 | 0.10 | | | 12/19 | 37/53 | 29.5 | 0.69 | 15/26 | 32/45 | 35.9 | 0.48 |
| 0.10 | 0.10 | 0.70 | 0.90 | 6/9 | 22/28 | 17.8 | 0.54 | 11/16 | 20/25 | 20.1 | 0.55 |
| 0.05 | 0.20 | | | 4/6 | 22/27 | 14.8 | 0.58 | 19/23 | 21/26 | 23.2 | 0.95 |
| 0.05 | 0.10 | | | 11/15 | 29/36 | 21.2 | 0.70 | 13/18 | 26/32 | 22.7 | 0.67 |

Table 2 Designs for $p_1 - p_0 = 0.15^a$

| α | β | p_0 | p_1 | Optimal Design | | | | Minimax Design | | | |
|----------|---------|-------|-------|------------------------------|------------|--------------------------|------|------------------------------|------------|--------------------------|------|
| | | | | Reject Drug if Response Rate | | EN(p_0) PET(p_0) | | Reject Drug if Response Rate | | EN(p_0) PET(p_0) | |
| | | | | $\leq r_1/n_1$ | $\leq r/n$ | | | $\leq r_1/n_1$ | $\leq r/n$ | | |
| 0.10 | 0.10 | 0.05 | 0.20 | 0/12 | 3/37 | 23.5 | 0.54 | 0/18 | 3/32 | 26.4 | 0.40 |
| 0.05 | 0.20 | | | 0/10 | 3/29 | 17.6 | 0.60 | 0/13 | 3/27 | 19.8 | 0.51 |
| 0.05 | 0.10 | | | 1/21 | 4/41 | 26.7 | 0.72 | 1/29 | 4/38 | 32.9 | 0.57 |
| 0.10 | 0.10 | 0.10 | 0.25 | 2/21 | 7/50 | 31.2 | 0.65 | 2/27 | 6/40 | 33.7 | 0.48 |
| 0.05 | 0.20 | | | 2/18 | 7/43 | 24.7 | 0.73 | 2/22 | 7/40 | 28.8 | 0.62 |
| 0.05 | 0.10 | | | 2/21 | 10/66 | 36.8 | 0.65 | 3/31 | 9/55 | 40.0 | 0.62 |
| 0.10 | 0.10 | 0.20 | 0.35 | 5/27 | 16/63 | 43.6 | 0.54 | 6/33 | 15/58 | 45.5 | 0.50 |
| 0.05 | 0.20 | | | 5/22 | 19/72 | 35.4 | 0.73 | 6/31 | 15/53 | 40.4 | 0.57 |
| 0.05 | 0.10 | | | 8/37 | 22/83 | 51.4 | 0.69 | 8/42 | 21/77 | 58.4 | 0.53 |
| 0.10 | 0.10 | 0.30 | 0.45 | 9/30 | 29/82 | 51.4 | 0.59 | 16/50 | 25/69 | 56.0 | 0.68 |
| 0.05 | 0.20 | | | 9/27 | 30/81 | 41.7 | 0.73 | 16/46 | 25/65 | 49.6 | 0.81 |
| 0.05 | 0.10 | | | 13/40 | 40/110 | 60.8 | 0.70 | 27/77 | 33/88 | 78.5 | 0.86 |
| 0.10 | 0.10 | 0.40 | 0.55 | 16/38 | 40/88 | 54.5 | 0.67 | 18/45 | 34/73 | 57.2 | 0.56 |
| 0.05 | 0.20 | | | 11/26 | 40/84 | 44.9 | 0.67 | 28/59 | 34/70 | 60.1 | 0.90 |
| 0.05 | 0.10 | | | 19/45 | 49/104 | 64.0 | 0.68 | 24/62 | 45/94 | 78.9 | 0.47 |
| 0.10 | 0.10 | 0.50 | 0.65 | 18/35 | 47/84 | 53.0 | 0.63 | 19/40 | 41/72 | 58.0 | 0.44 |
| 0.05 | 0.20 | | | 15/28 | 48/83 | 43.7 | 0.71 | 39/66 | 40/68 | 66.1 | 0.95 |
| 0.05 | 0.10 | | | 22/42 | 60/105 | 62.3 | 0.68 | 28/57 | 54/93 | 75.0 | 0.50 |
| 0.10 | 0.10 | 0.60 | 0.75 | 21/34 | 47/71 | 47.1 | 0.65 | 25/43 | 43/64 | 54.4 | 0.46 |
| 0.05 | 0.20 | | | 17/27 | 46/67 | 39.4 | 0.69 | 18/30 | 43/62 | 43.8 | 0.57 |
| 0.05 | 0.10 | | | 21/34 | 64/95 | 55.6 | 0.65 | 48/72 | 57/84 | 73.2 | 0.90 |
| 0.10 | 0.10 | 0.70 | 0.85 | 14/20 | 45/59 | 36.2 | 0.58 | 15/22 | 40/52 | 36.8 | 0.51 |
| 0.05 | 0.20 | | | 14/19 | 46/59 | 30.3 | 0.72 | 16/23 | 39/49 | 34.4 | 0.56 |
| 0.05 | 0.10 | | | 18/25 | 61/79 | 43.4 | 0.66 | 33/44 | 53/68 | 48.5 | 0.81 |
| 0.10 | 0.10 | 0.80 | 0.95 | 5/7 | 27/31 | 20.8 | 0.42 | 5/7 | 27/31 | 20.8 | 0.42 |
| 0.05 | 0.20 | | | 7/9 | 26/29 | 17.7 | 0.56 | 7/9 | 26/29 | 17.7 | 0.56 |
| 0.05 | 0.10 | | | 16/19 | 37/42 | 24.4 | 0.76 | 31/35 | 35/40 | 35.3 | 0.94 |

^aFor each value of (p_0, p_1) , designs are given for three sets of error probabilities (α, β) . The first, second and third rows correspond to error probability limits $(0.10, 0.10)$, $(0.05, 0.20)$, and $(0.05, 0.10)$ respectively. For each design, $EN(p_0)$ and $PET(p_0)$ denote the expected sample size and the probability of early termination when the true response probability is p_0 .

4

Quel design choisir ?

Dans l'hypothèse d'une molécule inefficace, le design Optimal de la méthode de Simon, pour lequel la probabilité de terminer l'étude de façon précoce est importante, permet de minimiser le nombre moyen de sujets requis.

En revanche, dans le cas d'une **maladie rare**, impliquant un **rythme d'inclusion très faible**, le **design Minimax** de Simon est **préférable** car la **taille maximale de l'échantillon** est alors **inférieure à celle du design Optimal**. Ainsi, dans le cas d'une étude avec $p_0 = 0,10$, $p_1 = 0,30$, $\alpha = \beta = 0,10$, le design Optimal de Simon correspond à une taille maximale d'échantillon de 35 sujets, alors que le design Minimax

correspond à seulement 25 sujets.

5

Conclusion

Parmi les plans de phase II à un seul bras, la méthode de Gehan en 2 étapes, très utilisée car d'un maniement facile, a l'in-

LA MÉTHODE DE SIMON EN PRATIQUE

convénient d'exposer un grand nombre de sujets à la 2^e étape si le produit se révèle inefficace. Les plans de Fleming et de Simon, bien que d'une utilisation moins aisée tentent de remédier à cet inconvénient.

La méthode de Fleming, à 2 ou 3 étapes, fixe dès le départ la taille de l'échantillon nécessaire à partir des valeurs des paramètres α , β , p_0 et p_1 . Elle permet de déterminer précocement, sur un effectif réduit, si un produit doit être rejeté ou au contraire est admissible à un développement en phase III. Néanmoins, les taux de réponse retenus pour conclure à l'efficacité ou à l'inefficacité de la molécule sont « extrêmes », très nettement supérieurs à p_1 ou très nettement inférieurs à p_0 .

A l'inverse de la méthode de Fleming, la taille de l'échantillon dans la méthode de Simon à 2 étapes est déterminée pour répondre à des critères de minimisation et ne peut être fixée à l'avance. Le design Optimal de Simon **permet de terminer l'essai précocement et de minimiser le nombre moyen de sujets nécessaires**, dans l'hypothèse d'un produit efficace. Le design Minimax de Simon qui minimi-

se le nombre maximum de sujets nécessaires peut être plus attractif que le design Optimal dans le cas où la pathologie concernée est rare et le taux de succès peu élevé.

6

Bibliographie

- 1. Armitage P., Mc Pherson C.K., Rowe B.C.**
Repeated significance tests on accumulating data
Journal of the Royal Statistical Society 1969; Series A 132: 235-244
- 2. Kramar A., Potvin D., Hill C.**
Plans expérimentaux pour l'inclusion de patients dans les essais de phase II
Revue Épidémiologie et Santé Publique 1996 ; 44 : 364-371
- 3. Lee Y.J., Staquet M., Simon R., Cantano R., Muggia F.**
Two-stage plans for patient accrual in phase II cancer clinical trials
Cancer Treatment Reports 1979; Vol 63: 1721-1726
- 4. Simon R.**
Design, analysis and reporting of cancer trials
In: Biopharmaceutical Statistics for Drug Development (Peace, KE, Ed.) New-York: Marcel Dekker, 1987
- 5. Simon R.**
Optimal two-stage designs for phase II clinical trials
Control Clin Trials 1989 Mar; 10(1): 1-10
Elsevier Science Publishing Co., Inc. 1989
- 6. Simon R., Wittes R.E., Ellenberg S.S.**
Randomized phase II clinical trials
Cancer Treatment Reports December 1985; Vol 69 N° 12
- 7. Sylvester R., Jand S., Aquet M.J.**
An application of decision theory to phase II clinical trials in cancer
In recent advances in cancer treatment
H.J. Tagnon and M.J. Stagnet (eds) 1-11 – New-York: Raven Press
- 8. Wittes R.E., Marsoni S., Simon R., Leyland-Jones B.**
The phase II trial
Cancer Treatment Reports November 1985; Vol 69 N° 11: 1235-1239

