



Design adaptatif de phase II/III seamless avec prise de décision bayésienne

Louise Bonnassieux

► To cite this version:

Louise Bonnassieux. Design adaptatif de phase II/III seamless avec prise de décision bayésienne. Annales de l'ISUP, 2017, 61 (3). <hal-03605993>

HAL Id: hal-03605993

<https://hal.science/hal-03605993v1>

Submitted on 11 Mar 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

ANNALES DE L'I.S.U.P.

Publications de l'Institut de Statistique de l'Université de Paris

Sommaire

D. BOSQ

Préface

L. BONNASSIEUX

Design adaptatif de phase II/III seamless

avec prise de décision bayésienne 1

BnF
DEP

Volume 61 – Fascicule 3 – 2017– PARIS

02

8° V 64396

ANNALES DE L'ISUP

Publications de l'Institut de Statistique de l'Université de Paris

Sommaire

D. BOSQ	
Préface	
J. BONASSIEUX	
Design adaptatif de plans IVIT séquentiels	
avec prise de décision bayésienne	1



18-05-18 BNF DLP 1 32

Volume 61 - Fascicule 3 - 2017 - PARIS

18-05-18 BNF DLP 1 32

25 8^{me} V 61301

**Éditorial pour le volume des Annales de l'ISUP
consacré au Prix Georges Darmois**

Denis Bosq
LPSM, UMR 8001, Paris

Créé en 1992, le prix Georges Darmois a pour but de récompenser le meilleur mémoire de 3ème année de l'ISUP. Georges Darmois (1888-1960) est un éminent Statisticien français, un des premiers en France à y enseigner la Statistique. Il participe à l'essor de l'Institut de Statistique de l'Université de Paris (ISUP) créé en 1922, et devient directeur de cet institut à la suite d'Émile Borel. Il est élu à l'Académie des sciences en 1955. En 2012, Gilbert Saporta, alors président de l'association des anciens étudiants (AAE) de l'ISUP, et Christian Partrat, directeur de l'ISUP, décident de créer le prix Georges Darmois en son hommage.

Les Annales de l'ISUP furent également fondées par Georges Darmois en 1952 sous le nom originel des *Publications de l'ISUP*. En tant que rédacteur en chef actuel et ancien directeur de l'ISUP, j'ai choisi de lui rendre également hommage en publiant les mémoires récents s'étant vus décernés le prix. Ce fascicule des Annales est ainsi consacré au travail de Mme Louise Bonnassieux, filière Biostatistique, qui a remporté le prix en 2015 pour son mémoire intitulé *Design adaptatif de phase II/III seamless avec prise de décision bayésienne* suite à son stage effectué à l'Institut de Recherches Internationales Servier dans le pôle d'expertise *Méthodologie et Valorisation des Données*. Je suis particulièrement heureux de la publication de ce mémoire qui met en valeur l'importance des techniques inférentielles et bayésiennes pour une application aussi sensible que l'oncologie. Ces techniques peuvent ainsi permettre de planifier des essais avec moins de patients tout en atteignant la même précision qu'avec un design classique.

Pour conclure, je tiens à remercier l'équipe de l'Association des Anciens Élèves de l'ISUP (AAE-ISUP) pour sa mobilisation dans la vie de l'ISUP, avec notamment l'organisation du Forum Emploi, événement au cours duquel elle décerne, avec l'ISUP, le prix Georges Darmois. Merci donc tout particulièrement à son président, Sébastien Bachelier, et à son assistante de direction, Suzanne Molinier, pour l'aide qu'ils m'ont apportée dans la recherche des lauréat(e)s des années précédentes.

Denis Bosq,
Éditeur des Annales de l'ISUP

Éditeur des Annales de l'ISUP
Denis Boad

années précédentes

Mais, pour l'aire d'aire m'ont apportée dans la recherche des statistiques des
son président, Sébastien Bachet, et à son assistant de direction, Suzanne
démarré avec l'ISUP, le prix Georges Darmois. Mais, dans tout particulièrement à
notamment l'organisation du Forum Emploi, événement au cours duquel elle
Élèves de l'ISUP (AAE-ISUP) pour sa mobilisation dans la vie de l'ISUP, avec
Pour conclure, je tiens à remercier l'équipe de l'Association des Anciens

présenté de planifier des essais avec moins de collets tout en atteignant le même
précision qu'avec un design classique.
même qui met en valeur l'importance des techniques informatiques et expérimentales
Validation des Données. Je suis particulièrement heureux de la production de ce
Recherches Internationales. Servir dans le rôle d'expert, Méthodologie et
séminaires avec pour thème la diversité humaine et son impact sur la qualité de
remporté le prix en 2015 pour son mémoire intitulé Design adaptatif de classe III
est ainsi consacré au travail de Mme Louise Bonnet-Darmois, élève de l'ISUP, qui a
publié les données récentes s'étant vu décerner le prix. Ce fascicule des Annales
actuel et ancien directeur de l'ISUP, l'air d'air de la recherche également hommage en
1982 sous le nom original des Publications de l'ISUP. En tant que rédacteur en chef
Les Annales de l'ISUP l'ont également fondées par Georges Darmois en

l'ISUP, décide de créer le prix Georges Darmois en son honneur.
l'association des anciens étudiants (AAE) de l'ISUP et Christian Fauriol, directeur de
l'Académie des sciences en 1986. En 2012, Gilbert Gauthier, alors président de
1922, et devient directeur de cet Institut à la suite d'Émile Borel. Il est élu à
partage à l'essor de l'Institut de Statistique de l'Université de Paris (ISUP) créé en
Statistique française, un des premiers en France à y enseigner la Statistique. Il
mémoire de 3ème année de l'ISUP, Georges Darmois (1882-1960) est un éminent
Créé en 1982, le prix Georges Darmois a pour but de récompenser le meilleur

L'ISUP, 1001, Paris
Denis Boad

consacré au prix Georges Darmois
Éditeur des Annales de l'ISUP



ISUP

PROMOTION 2014

Mémoire présenté devant

**L'Institut de Statistique
de l'Université Pierre et Marie Curie**

Pour l'obtention du

**Diplôme de Statisticien
Mention Biostatistique**

Par Melle Bonnassieux Louise

Sujet : design adaptatif de phase II/III seamless avec prise de décision bayésienne

Responsable du stage : Frédéric Dubois

Lieu du stage : Institut de Recherches Internationales Servier
Pôle d'expertise Méthodologie et Valorisation des Données

Invité(s) :

CONFIDENTIEL





UPMC
UNIVERSITÉ PARIS
MARTIN LUTHER KING

ISUP

PROMOTION 2014

Mémoire présenté devant

L'Institut de Statistique

de l'Université Pierre et Marie Curie

Pour l'obtention du

Diplôme de Statisticien

Mention Biostatistique

Par Mlle Bonnefleur Lucile

Sujet : design adaptatif de phase II/III randomisés avec prise de décision bayésienne

Responsable du stage : Frédéric Dubois

Lieu du stage : Institut de Recherche Internationale Génier
Pôle d'Expertise Méthodologie et Validation des Données

Intitulé :

☐ CONFIDENTIEL

Remerciements

Je tiens tout d'abord à remercier Christine Gabarroca pour m'avoir accueilli au sein de son unité et proposé un sujet fort intéressant.

Je remercie ensuite vivement Stéphanie Malassé, mon maître de Stage, puis Frédéric Dubois, qui a pris le relais, de m'avoir fait confiance dans la réalisation de ce projet, leur disponibilité et leur suivi pédagogique et humain tout au long de mon stage.

Merci aussi à l'ensemble de l'équipe ainsi que mes collègues de bureau pour leur accueil chaleureux et tous les moments partagés.

Enfin, je remercie également les enseignants et l'administration de l'ISUP pour la précieuse formation qui y est dispensée et leur accompagnement en continu.

Paris, Octobre 2014

Table des matières

Liste des figures et des tables.....	4
Présentation du stage.....	5
I- Cadre du stage : Les Laboratoires SERVIER.....	5
1- Le groupe SERVIER en quelques chiffres	5
2- La recherche et le développement.....	5
3- Les différents axes thérapeutiques	5
4- Pôle d'expertise Méthodologie et Valorisation des Données.....	6
II- Essais thérapeutiques.....	6
1- Le développement préclinique.....	7
2- Les études de phase I	7
3- Les études de phase II	7
4- Les études de phase III	7
5- L'Autorisation de Mise sur le Marché (AMM)	8
6- Les études de phase IV	8
III- Présentation du sujet.....	8
Le paradigme bayésien.....	10
I- Notions fondamentales.....	10
1- Le théorème de Bayes.....	10
2- La distribution a priori.....	10
3- La distribution a posteriori.....	11
4- La probabilité prédictive.....	13
5- L'approche bayésienne vs l'approche fréquentiste.....	13
II- L'analyse bayésienne dans un essai de phase II en Oncologie.....	14
1- Les essais de phase II en oncologie	14
2- La taille de l'échantillon	15
a- Détermination du nombre de sujets pour atteindre un degré de précision spécifié.....	15
b- Détermination du nombre de sujet qui satisfait un critère de puissance.....	17
3- Les analyses intermédiaires	21
4- Les probabilités prédictives.....	25

Les designs adaptatifs de phase II/III seamless.....	27
I- Les outils fréquentistes.....	29
1- Le principe de fermeture.....	29
2- La procédure de test adaptatif.....	30
II- Calcul de la puissance prédictive.....	32
1- Le contexte	32
a- Test de non-infériorité.....	32
b- Test de supériorité.....	32
c- Le jeu de donnée	32
2- La puissance prédictive.....	34
a- Le modèle paramétrique.....	34
b- Calcul de la puissance conditionnelle.....	36
c- Calcul de la puissance prédictive.....	36
Exemples.....	38
I- Application.....	38
1- Le contexte	38
2- Le calcul de la puissance prédictive	38
a- Distribution a priori.....	39
b- Jeux de données à l'analyse intermédiaire.....	39
c- Simulation des données manquantes	40
d- Résultats	41
II- Perspectives : l'enrichissement de population	43
Conclusion.....	46
Annexes.....	47
I. Annexe 1 : Calcul de la distribution a posteriori.....	47
II. Annexe 2 : Algorithme pour la table 3.....	48
III. Annexe 3 : Détail de la chaîne de Markov.....	49
IV. Annexe 4 : Les analyses intermédiaires.....	51
V. Annexe 5 : Calcul de la probabilité prédictive.....	54
VI. Annexe 6 : Calcul des taux de hasard a posteriori.....	55
VII. Annexe 7 : Code R des tables 7,9, 10 et 11	55
Bibliographie.....	67

Liste des tables et des figures

Figure 1 : organisation du pôle méthodologie et valorisation des données.....	6
Figure 2 : développement classique d'un essai thérapeutique.....	7
Figure 3 : densité a priori, vraisemblance et a densité a posteriori d'une distribution $\beta(1,1)$ (non informative) et $\beta(8,5)$ (informative).....	12
Figure 4 : bornes de l'intervalle de crédibilité et densité a posteriori d'une loi beta telle que $\mu_E = 0.30$, $p_I = 0.9$ et $\Delta_I = 0.2$	17
Figure 5 : probabilité de déclarer le traitement expérimental (E) prometteur en fonction de la taille de l'échantillon M, avec la vraie probabilité de succès de E allant de 0.50 à 0.70, $\mu_s = 0.50$, $\delta_0 = 0.20$, $W_{s,90} = 0.30$ et $(p_u, p_l) = (0.05, 0.95)$	21
Figure 6 : scénarios possible à la première analyse intermédiaire.....	22
Figure 7 : scénarios possible à la deuxième analyse intermédiaire.....	23
Figure 8 : scénarios possible l'analyse finale page.....	24
Figure 9 : évolution du taux de réponse du traitement expérimental.....	25
Figure 10 : schéma classique d'un essai clinique.....	27
Figure 11 : phase II/III seamless sélection de dose.....	28
Figure 12 : phase II/III seamless enrichissement de population.....	28
Figure 13 : principe de fermeture pour 3 hypothèses nulles.....	30
Figure 14 : design adaptatif en deux étapes.....	31
Figure 15 : schéma de suivi des patients.....	34
Figure 16 : calcul de la puissance prédictive.....	42
Figure 17 : schéma population.....	44
Figure 18 : Détails de la chaîne de Markov permettant de calculer le nombre de sujet nécessaire.....	49
Table 1 : Exemple de distribution a priori non-informative.....	11
Table 2 : Exemple de lois conjuguées.....	11
Table 3 : Table 3 : taille d'échantillon nécessaire pour assurer un degré de précision spécifié.....	16
Table 4 : nombre de sujet nécessaire pour $\mu_s = 0.2$	20
Table 5 : nombre de sujet nécessaire $\mu_s = 0.5$	20
Table 6 : erreurs dans les tests d'hypothèses.....	29
Table 7 : paramètres a priori des taux de hasard pour chaque intervalle de temps et la fonction de répartition.....	39
Table 8 : nombre d'événement dans chaque groupe de traitement à l'analyse intermédiaire.....	40
Table 9 : distribution a posteriori des taux de hasard.....	41
Table 10 : taux d'événement à 6 mois a posteriori lors de l'analyse intermédiaire, puissance prédictive pour les tests de non infériorité et de supériorité.....	42
Table 11 : caractéristiques d'un design adaptatif de phase II/III seamless utilisant la PP pour sélectionner la meilleure dose ou arrêter l'essai pour futilité.....	43

Présentation du stage

I- Cadre du stage : Les Laboratoires SERVIER

1- Le groupe SERVIER en quelques chiffres

Créé à Orléans en 1954 par le docteur Jacques Servier (1922-2014) avec 9 collaborateurs, l'institut SERVIER est aujourd'hui le premier groupe pharmaceutique indépendant français, et le deuxième groupe pharmaceutique français au niveau mondial.

Depuis presque 60 ans, les laboratoires SERVIER poursuivent leur développement, aussi bien en France qu'à l'international, où sont consommés 90% de leurs médicaments. Le groupe est présent dans plus de 140 pays, représentant plus de 21 000 collaborateurs dont 3 000 en recherche et développement et 5 000 en France.

La mission des laboratoires SERVIER est de mettre au point de nouveaux médicaments. Pour ce faire, leurs activités concernent toutes les étapes de la vie de ce dernier : recherche, production, information des médecins en France et à l'international ...

2- La recherche et le développement

Tous les médicaments mis sur le marché par les laboratoires SERVIER sont issus de sa propre recherche, et avec 30 médicaments innovants en 30 ans reconnus par tous les praticiens et vendus dans le monde entier, le laboratoire a démontré son savoir-faire et son expertise pharmacologique.

Près de 27% d'un chiffre d'affaires annuel du groupe s'élevant à 4 200 milliard d'euros est consacré à la recherche et au développement avec près de 3 000 collaborateurs dans ce secteur (R&D) et 19 Centres Internationaux Thérapeutiques (CIRT).

3- Les différents axes thérapeutiques

Les laboratoires SERVIER sont titulaires de plus de 27 000 brevets déposés pour plus de 50 000 molécules synthétisées. Le groupe s'assure un pôle de recherche dynamique et important dont les principales aires thérapeutiques sont :

- la cardiologie
- la neuropsychiatrie
- le métabolisme
- la rhumatologie
- l'hypertension artérielle
- l'oncologie

Par ailleurs, en 1996, face à l'émergence du marché des génériques, le Dr Servier a créé la filiale *Biogaran*. Avec près de 25% de parts de marché, la filiale occupe aujourd'hui une place majeure sur ce marché.

4- Pôle d'expertise Méthodologie et Valorisation des Données

Le pôle d'expertise *Méthodologie et Valorisation des Données* a les missions suivantes :

- Optimiser les méthodologies relatives aux études cliniques et pharmaco-épidémiologiques.
- Participer à la conception d'études cliniques et pharmaco-épidémiologiques, réaliser l'analyse des données issues de ces études en se basant sur des méthodes statistiques approuvées par les autorités de santé, procéder à l'interprétation des résultats et à la production de rapport d'études cliniques et d'analyses intégrées.
- Évaluer l'acceptabilité des produits du laboratoire dans une dimension populationnelle, détecter les signaux et évaluer les risques.
- Développer une expertise biostatistique et pharmaco-épidémiologique.

Ce pôle d'expertise est constitué de huit départements dirigés par Maylis Coste :

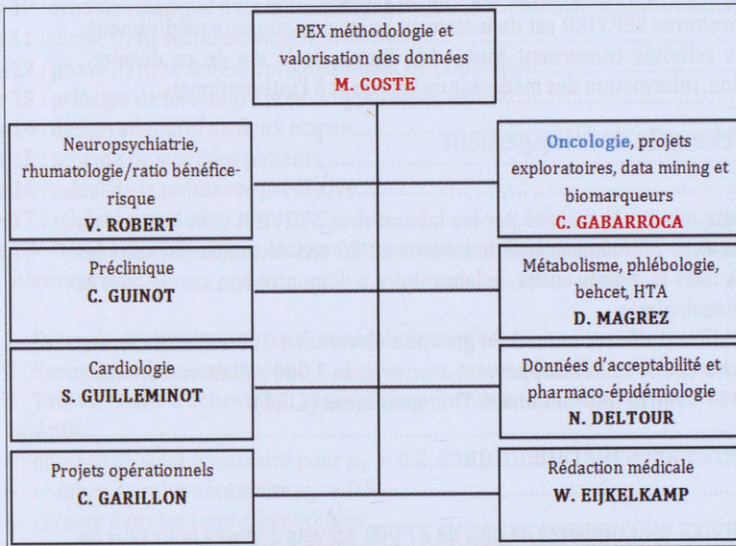


Figure 1 : organisation du pôle méthodologie et valorisation des données

II- Essais thérapeutiques

Entre la découverte de la molécule et la mise sur le marché d'un nouveau médicament, il s'écoule en général entre 7 et 12 ans, pendant lesquels le médicament est soumis à un certain nombre de tests démontrant son efficacité. Sur 10 000 molécules potentielles, seulement 20 parviendront au stade des phases cliniques, et uniquement une atteindra le stade de l'Autorisation de Mise sur le Marché (AMM).

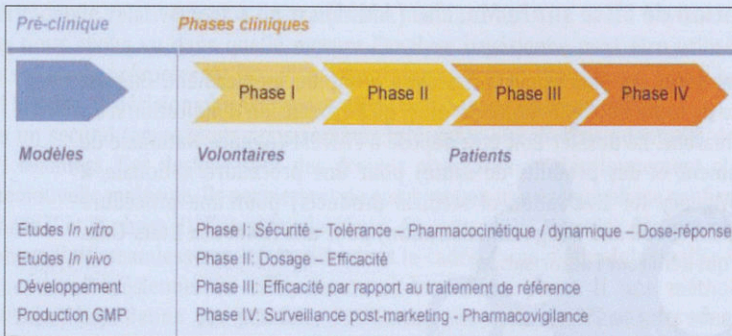


Figure 2 : développement classique d'un essai thérapeutique

1- Le développement préclinique

Dans un premier temps, des études sont réalisées chez l'animal, afin d'étudier les propriétés pharmacocinétiques, pharmacodynamiques et toxicologiques de la molécule. C'est au cours de cette étape que seront étudiés entre autres, les effets secondaires, la toxicité, les conséquences sur la reproduction et sur le développement éventuel de tumeurs. Les données recueillies sont indispensables au bon déroulement de l'étude car, au-delà de la phase préclinique, l'administration se fait sur l'homme.

2- Les études de phase I

Les études ont pour objectif de définir la tolérance humaine du médicament, la dose maximale tolérée et, si possible, la posologie entraînant les premiers effets indésirables ainsi que celle entraînant les premiers effets pharmacodynamiques souhaités. Ces études sont le plus souvent réalisées chez quelques dizaines de sujets volontaires sains, sauf en oncologie où les patients sont tous des volontaires malades, en fin de vie, pour lesquels tous les autres traitements ont échoué.

3- Les études de phase II

Les études de phase II sont des études exploratoires qui visent à déterminer l'efficacité pharmacologique du médicament. Elles permettent de recueillir des données de tolérance de pharmacocinétique et de relation dose-effet, mais aussi de réaliser l'étude au sein de sous-groupes de populations fragiles (personnes âgées, insuffisants rénaux, hépatiques...).

4- Les études de phase III

Les études de phase III sont des études confirmatoires qui permettent de prouver l'efficacité thérapeutique du médicament dans les différentes indications revendiquées. L'effet thérapeutique du produit est étudié sur un groupe de patients homogène souffrant de la maladie à traiter. Il est mesuré sur un paramètre précis et comparé par rapport à un placebo ou à un traitement de référence. Chaque essai (quelques centaines à quelques milliers de personnes) est organisé selon un protocole précis : nombre de patients à inclure calculé à l'avance, essai conduit en double aveugle.

5- L'Autorisation de Mise sur le Marché (AMM)

L'AMM est l'une des dates les plus importantes dans la vie du médicament. Elle est nécessaire à la vente de toute nouvelle molécule, ainsi qu'à l'extension d'indication d'un produit déjà sur le marché. Le dossier doit être déposé à l'ANSM (Agence Nationale de Sécurité du Médicament et des produits de santé) pour une procédure nationale, à l'EMA (European Agency for Evaluation of Medical products) pour une procédure Européenne et à la FDA (Food and Drug Administration) pour une AMM aux Etats-Unis. Ce sont ces agences qui délivrent l'autorisation.

6- Les études de phase IV

Les études de phase IV ou de pharmacovigilance, ont lieu après la mise sur le marché du médicament. Elles permettent de mieux évaluer l'utilité réelle et la tolérance du nouveau médicament au sein d'une population plus large et plus représentative dans des conditions usuelles de prescription et de suivi.

III- Présentation du sujet

En oncologie, l'émergence de nouvelles thérapies, comme par exemple les thérapies ciblées ou les immunothérapies, confronte les industriels à des choix de doses de plus en plus compliquées ou des choix de sous populations systématiques. Développer une thérapie ciblant une ou plusieurs voies biochimiques suggère qu'une sous-population spécifique de patients est plus susceptible d'obtenir une réponse à ce traitement ou encore de sélectionner des doses de traitement pertinentes. La possibilité de sélectionner ces patients ou ces doses est cruciale pour le développement d'une telle thérapie. Mais cette identification dans un plan de développement classique peut être longue et coûteuse.

Lors de ce stage nous nous sommes alors demandé dans quelle mesure les designs adaptatifs de phase II/III peuvent répondre à nos exigences et également quels apports peuvent avoir les statistiques bayésiennes dans notre prise de décision.

Selon la règle de Bayes, la probabilité d'un événement n'est pas uniquement déterminée par les résultats expérimentaux nouveaux, mais aussi par des résultats antérieurs. En effet, l'analyse bayésienne prend en compte toute l'information préexistante et n'ignore ni les informations, ni le manque d'informations pour modéliser la distribution de notre paramètre et utilise toute l'information disponible pour mettre à jour notre connaissance/incertitude du paramètre. Ces méthodes correspondent bien au raisonnement par inférence utilisé en sciences naturelles et dans la recherche médicale. Dans un premier temps, et pour me familiariser avec l'analyse bayésienne, qui était une notion nouvelle pour moi, nous avons travaillé dans le cadre d'un essai de phase II en oncologie. Les essais de phase II jouent un rôle important dans le processus de développement d'un nouveau traitement. En oncologie, la phase II est une phase d'apprentissage où l'objectif est d'évaluer l'efficacité et la tolérance du traitement expérimental. Ainsi, seules les molécules ayant montré une efficacité anti-tumorale suffisante et une tolérance acceptable poursuivront leur développement dans des essais de phase III. Pour définir cette étude de phase II, nous avons commencé par déterminer le nombre de patients nécessaire mettant en évidence l'efficacité d'un traitement

expérimentale relativement à un traitement standard dans un cadre bayésien. Par la suite, nous avons vu dans quelle mesure l'analyse bayésienne peut être utilisée pour arrêter l'étude précocement lors d'analyses intermédiaires. Dans ce cas, l'outil bayésien rend la prise de décision plus transparente qu'une p-value en fréquentiste.

Dans un second temps, nous nous sommes intéressés aux designs adaptatifs de phase II/III seamless. Ces designs sont des designs alternatifs au développement classique d'une nouvelle molécule. Ils permettent de combiner en une unique phase confirmatoire la phase II et la phase III d'un essai classique. Au cours de ce stage les designs adaptatifs de phase II/III seamless ont été étudiés dans le cadre d'une méthodologie dite hybride fréquentiste/bayésienne. En effet, même si lors d'une phase II une méthodologie totalement bayésienne est possible, l'utilisation de cette dernière pour des essais confirmatoires n'est pas encore reconnue par les agences réglementaires. Les outils fréquentistes sont alors utilisés pour l'analyse finale et le contrôle de l'erreur de type I tandis que les principes bayésiens sont utilisés pour la prise de décision lors des analyses intermédiaires. Nous avons appliqué cette méthodologie statistique dans le contexte de recherche de dose et nous avons étendu ces résultats pour la sélection de population.

Le paradigme Bayésien

I- Notions fondamentales

1- Le théorème de Bayes

Notations

Soient :

- $X = (X_1, \dots, X_n)$ le n-uplet de variables aléatoires définies sur l'espace probabilisé $(\Omega, \mathcal{A}, P_\theta)$ et \mathbb{X} l'espace des réalisations de X
- $x = (x_1, \dots, x_n)$ les réalisations de X
- θ le vecteur des paramètres inconnus définis sur l'espace Θ
- $\pi(\theta)$ la distribution a priori de θ et $\pi(\theta|x)$ la distribution a posteriori de θ réactualisé en fonction des observations
- $f(x|\theta)$ la fonction de vraisemblance de x à θ fixé

Définition 1

Nous appelons **modèle statistique bayésien**, la donnée d'un modèle statistique paramétré $(\Omega, \mathcal{A}, P_\theta, \theta \in \Theta)$ avec $f(x|\theta)$ densité de P_θ et d'une loi $\pi(\theta)$ sur le paramètre.

Théorème 1- Théorème de Bayes

Avec les notations précédentes, si P_θ est absolument continue par rapport à une mesure ν pour tout $\theta \in \Theta$, alors la loi conditionnelle de θ sachant $X = x$ est absolument continue par rapport à une mesure μ pour presque tout x . Sa densité est :

$$\pi(\theta|x) = \frac{f(x|\theta) \cdot \pi(\theta)}{m(x)}$$

Où $m(x)$ est la distribution marginale de X définit par :

$$m(x) = \int f(x|\theta) \cdot \pi(\theta) d_\theta$$

Ainsi la transformation de la connaissance a priori en une connaissance a posteriori, en fonction des observations disponibles, se fait grâce au théorème de Bayes :

$$\pi(\theta|x) \propto \pi(\theta) \times f(x|\theta)$$

a posteriori \propto a priori \times vraisemblance

2- La distribution a priori

Comme nous avons pu le voir grâce à la formule de Bayes, la distribution a posteriori se déduit directement de la distribution a priori du paramètre et des observations disponibles. Le choix de la distribution a priori est de ce fait une étape fondamentale

dans l'analyse bayésienne. Ce choix est basé sur la connaissance que l'on a sur le paramètre. En effet dans le cadre d'un essai clinique, la bibliographie et les phases antérieures apportent de l'information. Il peut également être basé sur le manque d'information disponible. Deux grandes familles de distribution se distinguent alors pour le paramètre :

1- La distribution informative :

Dans ce cas, nous avons une bonne connaissance a priori du paramètre, notamment grâce aux informations extérieurs à l'étude en cours ou à la bibliographie. Les informations disponibles, via la vraisemblance, permettront d'affiner la précision sur la distribution du paramètre.

2- La distribution non-informative :

Dans ce cas, nous n'avons que très peu de connaissance a priori sur le paramètre. Pour modéliser ce manque d'informations, plusieurs distributions non-informatives existent selon la nature du paramètre (continu, binaire...).

Distribution a priori du paramètre	Forme non informative généralement utilisée
Distribution beta	$\mathcal{U}([0,1]) = \beta(1,1)$ (a priori uniforme) $\beta(1/2, 1/2)$ (a priori de Jeffrey)...
Distribution gamma	$\mathcal{G}(0,1), \mathcal{G}(1/2, 0)$
Distribution normale	$\mathcal{N}(\mu, \infty)$...

Table 1 : Exemple de distribution a priori non-informative

3- La distribution a posteriori

Une fois les paramètres de la distribution a priori fixés, la distribution a posteriori sachant les observations disponibles peut être calculée grâce au théorème de Bayes. Or, ce calcul n'est pas toujours évident mais peut être facilité lorsque les distributions a priori et a posteriori ont la même forme, nous parlons alors de distribution conjuguée.

Définition 2

Une famille \mathcal{F} de lois sur Θ est dite **conjuguée** si, pour tout π appartenant à cette famille, la loi $\pi(\theta|x)$ appartient également à celle-ci.

Le principal avantage des familles de distributions conjuguées est de simplifier les calculs des distributions a posteriori, en effet dans ce cas la distribution a posteriori se déduit directement de la distribution a priori.

$f(x \theta)$	$\pi(\theta)$	$\pi(\theta x)$
$\mathcal{B}(n, \theta)$	$\beta(\alpha, \gamma)$	$\beta(\alpha + x, \gamma + n - x)$
$\mathcal{G}(n, \theta)$	$\mathcal{G}(\alpha, \beta)$	$\mathcal{G}(\alpha + n, \beta + x)$
$\mathcal{N}(\theta, \sigma^2)$	$\mathcal{N}(\mu, \tau^2)$	$\mathcal{N}\left(\frac{x}{\sigma^2} + \frac{\mu}{\tau^2}, \left(\frac{1}{\sigma^2} + \frac{1}{\tau^2}\right)^{-1}\right)$

Table 2 : Exemple de lois conjuguées

Exemple : Illustrer l'impact d'un a priori informatif et non informatif sur une étude et calcul de la distribution a posteriori

Soit $X = (X_1, \dots, X_n)$ un échantillon indépendant et identiquement distribué suivant une loi de Bernoulli de paramètre θ . Dans le cadre de réponses binaires, le taux de réponse θ à un traitement est généralement bien modélisé par une distribution beta :

$$\pi(\theta) = \beta(a, b) \text{ avec } \mathbb{E}(\theta) = a/(a+b) \text{ et } \mathbb{V}(\theta) = a/(a+b)^2$$

Soient deux distributions a priori :

- $\theta \sim \beta(1,1)$ modélisant un a priori non informatif : correspond à une loi uniforme. La moyenne du taux de réponse est de 50% avec une variance de 0.25.
- $\theta \sim \beta(8,5)$ modélisant un a priori informatif : correspond à une moyenne du taux de réponse de 62% avec une variance de 0.05.

La vraisemblance est classiquement calculée comme étant le produit des densités :

$$L(x_i, \theta) = \prod_{i=1}^n \theta^{x_i} (1-\theta)^{1-x_i}, x_i \in \{0,1\}$$

$$L(x_i, \theta) = \theta^{\sum_{i=1}^n x_i} (1-\theta)^{n - \sum_{i=1}^n x_i} = \theta^s (1-\theta)^{n-s}$$

Nous en déduisons, grâce au théorème de Bayes, la loi a posteriori :

$$\pi(\theta|x) = \beta(a+n, b+s-n) \text{ (Démonstration annexe 1)}$$

Pour l'exemple, nous considérons que l'analyse est faite au bout de 20 patients recrutés ($n=20$). Nous observons 13 réponses au traitement ($s=13$), ce qui correspond à un taux de réponse de 65%.

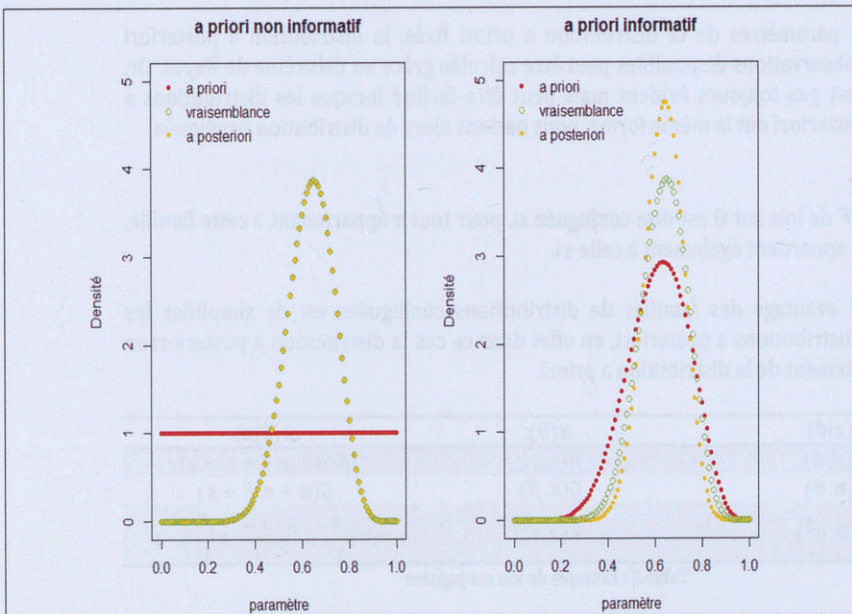


Figure 3 : densité a priori, vraisemblance et densité a posteriori d'une distribution $\beta(1,1)$ (non informative) et $\beta(8,5)$ (informatif)

Remarquons, sur le premier graphique de la figure 3, que pour un a priori au départ non informatif, la distribution a posteriori est totalement confondue avec la vraisemblance. Tandis que pour les mêmes données collectées, mais avec une distribution a priori informative, la distribution a posteriori est beaucoup plus précise, plus resserrée. En effet, la distribution du taux de réponse a une information supplémentaire à savoir les observations de X . Ainsi, l'incertitude sur le paramètre θ diminue.

Un potentiel gain d'avoir au début de l'essai un a priori informatif est que pour le même nombre de patients et le même taux de réponse, la variabilité diminue. Autrement dit, pour un niveau de précision souhaité, il faudrait moins de patients avec une distribution informative.

4- La probabilité prédictive

La probabilité prédictive permet de prédire grâce aux observations disponibles les observations futures. Ainsi, la meilleure prédiction possible est calculée compte tenu à la fois de notre savoir et de notre ignorance sur le paramètre d'intérêt.

Définition 3

Soit y une observation future, la **distribution prédictive** est définie de la façon suivante :

$$\mathbb{P}(y|x) = \int \pi(\theta|x) \mathbb{P}(y|\theta) d_{\theta}$$

Les probabilités prédictives peuvent être utilisées pour planifier l'essai en déterminant le nombre de patients à inclure mais peuvent aussi être très utiles lors d'analyses intermédiaires pour arrêter précocement l'étude, soit pour efficacité si la probabilité est élevée, soit pour futilité si la probabilité est faible.

Lors d'analyses intermédiaires, des données peuvent être manquantes, aussi les probabilités prédictives permettent de traiter ce cas.

5- L'approche bayésienne vs l'approche fréquentiste

L'idée fondamentale de l'approche bayésienne est que notre paramètre d'intérêt admet une distribution. C'est ce qui distingue l'analyse fréquentiste de la l'analyse bayésienne, en effet, en fréquentiste, le paramètre est considéré comme fixe et inconnu. Ici, notre paramètre admet une distribution qui sera mise à jour au cours de l'essai grâce aux observations accumulées.

En fréquentiste, nous parlons d'intervalle de confiance contenant la vraie valeur du paramètre. Cet intervalle est uniquement basé sur les données et ne prend pas en compte l'incertitude sur le paramètre d'intérêt. Cette incertitude est au contraire prise en compte avec l'intervalle de crédibilité en bayésien. En effet dans ce cas le paramètre est aléatoire et cet intervalle reflète notre force de conviction sur notre paramètre d'intérêt.

De plus, l'approche fréquentiste base ses prises de décisions sur la p-value (« le plus petit seuil de significativité pour lequel l'hypothèse nulle est acceptée »). Mais il est bien connu qu'en soi, « $p < 0,05$ » informe peu et mène fréquemment à des interprétations erronées. Les méthodes bayésiennes, quant à elles, permettent d'évaluer les effets

traitement sans imposer la recherche de valeurs significatives et apportent une réponse clinique facilement interprétable.

Les procédures bayésiennes permettent alors de répondre de façon naturelle à des questions telles que :

- « quelle est la probabilité de succès à la fin l'étude au vue des informations disponible ? »
- « quel est le nombre de patients à inclure pour pouvoir conclure positivement à la fin de l'essai ? »

Dans chacun des cas, le cadre bayésien, permet d'utiliser l'information disponible a priori pour estimer l'effet traitement. C'est un avantage majeur par rapport au fréquentiste, ici toute l'information disponible est utilisée dans l'analyse. L'information a priori est généralement basée sur la bibliographie, les résultats des phases antérieures et les résultats du début de l'essai en cours. Plus nous avons de l'information sur notre paramètre, plus la variabilité, i.e. la largeur de l'intervalle de crédibilité, est petite.

II- L'analyse bayésienne dans un essai de phase II en Oncologie

Dans ce chapitre, nous nous intéressons à l'utilisation de l'analyse bayésienne dans le cadre d'un d'essai de phase II en oncologie. Ce fut une étape importante du stage, qui m'a permis de mettre en perspective le paradigme bayésien.

1- Les essais de phase II en oncologie

L'objectif principal d'un essai clinique de phase II est d'obtenir les premières estimations de l'efficacité thérapeutique d'un nouveau traitement. Plus particulièrement, les essais de phase II standard sont composés de peu de patients, en général moins de 50. Ce sont des essais ouverts avec un seul bras de traitement et donc pas de randomisation des patients dans les groupes traités et contrôle. Plusieurs analyses intermédiaires sont réalisées afin d'arrêter précocement l'essai pour futilité. Enfin, à la fin de l'essai, des prises de décisions (GO/NO GO) sont effectuées pour aller en phase III.

Les outils fréquentistes ou bayésiens peuvent être utilisés dans ce contexte. Ces deux approches permettent de faire plusieurs analyses intermédiaires mais en fréquentistes il faudra tenir compte de l'erreur de type I tandis qu'en bayésien, puisque la notion de test n'existe pas, il n'y a pas de contrainte en terme de nombre d'analyses intermédiaires. Ainsi l'outil bayésien peut être utilisé de plusieurs manières :

- Pour déterminer le nombre de sujet minimal à inclure dans l'étude ou réévaluer le nombre de sujet au cours de l'étude lors d'analyses intermédiaires.
- Pour déterminer des règles d'arrêt de l'étude. Aux analyses intermédiaires, la probabilité que l'essai soit concluant à la fin de l'étude sachant ce qui est observé est calculée. Si cette probabilité est petite, l'essai est arrêté pour futilité ; Si cette

probabilité est grande, l'essai est arrêté pour efficacité ; Sinon l'essai continu. Les seuils de décision étant déterminés à l'avance.

- Pour prendre la décision d'aller ou non en phases III grâce aux informations de la phase II.

2- La taille de l'échantillon

Nous souhaitons déterminer le nombre de patients à inclure dans l'essai afin de mettre en évidence un effet traitement. Ceci dans un cadre bayésien, où l'efficacité du traitement est aléatoire. Dans le cadre du stage, nous avons repris et implémenté, grâce au logiciel R, les méthodes utilisées dans les articles [11,12].

En oncologie, les essais sont généralement composés de peu de patients, nous fixons alors une taille maximale d'échantillon à ne pas dépasser de 100 patients. Il est également important pour pouvoir conclure d'avoir une taille d'échantillon minimale, ici elle sera fixée à 10 patients.

Pour déterminer la taille de l'échantillon, deux méthodes ont été étudiées. Si l'essai ne s'arrête pas précocement et si le nombre maximal de patients fixé n'est pas atteint, le nombre de patients est calculé :

- tel que l'efficacité a posteriori du traitement expérimental atteigne un certain degré de précision.
- tel que l'efficacité a posteriori du traitement expérimental atteigne un critère de puissance.

Soit l'efficacité a priori d'un traitement expérimental θ_E basée sur des réponses binaires au traitement. Dans ce cas la distribution de θ_E est modélisée par une loi beta :

$$\theta_E \sim \beta(a_E, b_E)$$

$$\mathbb{E}(\theta_E) = \frac{a_E}{a_E + b_E} \text{ et } \mathbb{V}(\theta_E) = \frac{a_E}{(a_E + b_E)^2}$$

Soit M le nombre de patients à déterminer et X_M le nombre de réponses possibles pour M patients recrutés. Dans le cadre de loi conjuguée, la distribution a posteriori de θ_E conditionnellement aux observations est :

$$\theta_E | X_M \sim \beta(a_E + x_M, b_E + M - x_M) \text{ (Démonstration annexe 1)}$$

Ici la distribution de θ_E est considérée non-informative et ainsi les valeurs des paramètres pour la distribution a priori sont $a_E = b_E = 1$.

a- Détermination du nombre de sujets pour atteindre un degré de précision spécifié

Pour déterminer M la taille de l'échantillon telle que la distribution a posteriori du traitement expérimental $\theta_E | X_M = x_M$ atteigne un certain degré de précision, nous utilisons une méthode basée sur l'intervalle de crédibilité.

Soient $p_l \in [0.90, 0.99]$ une probabilité de couverture désirée et $\Delta_l = U - L$ une largeur d'intervalle désirée, avec U et L les quantiles à $0.5 + p_l/2$ et $0.5 - p_l/2$ respectivement de la distribution a posteriori du traitement expérimental.

La taille de l'échantillon M est choisie telle que :

$$\mathbb{P}(L < \theta_E < U | X_M = x_M) \geq p_l$$

Dans la formule précédente, la distribution du taux de réponse du traitement expérimental sachant $X_M = x_M$ réponses pour M patients est utilisée. Cette distribution correspond à la distribution a posteriori du taux de réponse du traitement expérimental. Cette distribution doit être calculée mais à ce moment de l'essai X_M est inconnu. Dans ce cas, les paramètres de la distribution a posteriori du traitement expérimental sont également inconnus. Or, nous souhaitons déterminer M permettant de détecter l'effet traitement, c'est-à-dire déterminer M permettant d'atteindre un taux de réponse ciblé. Soit μ_E le taux de réponse souhaité, X_M est calculé tel que :

$$\begin{aligned} \mathbb{E}(\theta_E | X_M) &= \frac{a_E + X_M}{a_E + b_E + M} = \mu_E \\ \Leftrightarrow X_M &= \mu_E(a_E + b_E + M) - a_E \\ \Leftrightarrow X_M &= \mu_E(1 + 1 + M) - 1 \end{aligned}$$

Nous en déduisons les résultats suivants (cf. *algorithme en annexe 2*) :

μ_E a posteriori	$\Delta_l = 0.2$		$\Delta_l = 0.25$	
	$p_l = 0.9$	$p_l = 0.95$	$p_l = 0.9$	$p_l = 0.95$
0.20	39	58	24	34
0.25	47	68	30	42
0.30	54	78	34	48
0.35	59	84	38	53
0.40	63	89	39	56
0.45	65	92	41	58

Table 3 : taille d'échantillon nécessaire pour assurer un degré de précision spécifié

Nous remarquons très logiquement que plus la largeur d'intervalle est petite et la probabilité de couverture est grande, plus le nombre de patients permettant de détecter l'effet traitement est grand.

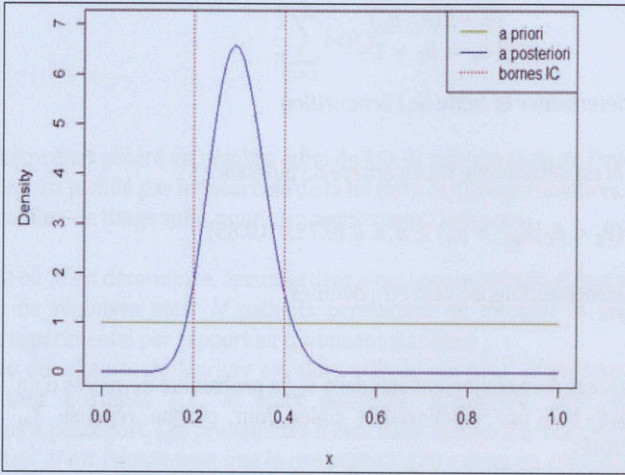


Figure 4 : bornes de l'intervalle de crédibilité et densité a posteriori d'une loi beta telle que $\mu_E = 0.30$, $p_I = 0.9$ et $\Delta_I = 0.2$

b- Détermination du nombre de sujet qui satisfait un critère de puissance

Pour déterminer M la taille de l'échantillon qui satisfait un critère de puissance le traitement expérimental est comparé à un traitement standard.

1- Informations a priori

a- Information sur le traitement standard

Soit un traitement standard de référence dont l'efficacité a déjà été prouvée. Ainsi, le taux de réponse de ce traitement et la largeur de l'intervalle de crédibilité autour de ce taux sont connus :

$$\begin{cases} \theta_S \sim \beta(a_S, b_S), a_S, b_S \text{ inconnus} \\ \mu_S \text{ taux de réponses connu et } W_{S,90} \text{ largeur de} \\ \text{l'intervalle de crédibilité à 90\% autour de } \mu_E \text{ connue} \end{cases}$$

Grâce à ces informations sur le taux de réponse, les paramètres de la distribution de θ_S sont déterminés en résolvant les équations suivantes :

$$\begin{cases} \mu_S = \frac{a_S}{a_S + b_S} \Leftrightarrow a_S = b_S \frac{\mu_S}{1 - \mu_S} \\ b_S = \min_{b_S} \{ ([q_{0.95} \left(\beta \left(b_S \frac{\mu_S}{1 - \mu_S}, b_S \right) \right) - q_{0.05} \left(\beta \left(b_S \frac{\mu_S}{1 - \mu_S}, b_S \right) \right)] - W_{S,90} \}^2 \} \end{cases}$$

b- information sur le traitement expérimental

L'a priori sur le traitement expérimental est non informatif :

$$\begin{cases} \theta_E \sim \beta(a_E, b_E) \\ a_E = b_E = 1 \end{cases}$$

2- Critère pour déterminer la taille de l'échantillon

La taille de l'échantillon M est déterminée via un critère de puissance :

$$\mathbb{P}(\mathbb{P}(\theta_S < \theta_E | X_M) > p_u) \geq \pi, \pi \in \{0.75, 0.8, 0.85\}$$

Avant de calculer M , plusieurs notions doivent être définies.

Définition 4

Soient la probabilité de succès d'un traitement standard θ_S , la probabilité de succès d'un traitement expérimental θ_E et δ_0 une amélioration ciblée. Pour chaque réponse X_M possible à M patients recrutés :

- E est déclaré **prometteur** par rapport à S si :

$$\mathbb{P}(\theta_S < \theta_E | X_M) = \iint \beta_{\theta_S}(a_S, b_S) \beta_{\theta_E}(a_E + x_M, b_E + M - x_M) \mathbb{I}_{\theta_S < \theta_E} d\theta_S d\theta_E > p_u,$$

$$p_u \in [0.9, 0.99]$$

C'est-à-dire qu'avec une certaine force de conviction, le traitement expérimental est supérieur au traitement standard.

- E est déclaré **non prometteur** par rapport à S si :

$$\mathbb{P}(\theta_S + \delta_0 < \theta_E | X_M) = \iint \beta_{\theta_S}(a_S, b_S) \beta_{\theta_E}(a_E + x_M, b_E + M - x_M) \mathbb{I}_{\theta_S + \delta_0 < \theta_E} d\theta_S d\theta_E < p_l,$$

$$p_l \in [0.01, 0.1]$$

C'est-à-dire qu'avec une certaine force de conviction, le traitement expérimental n'est pas strictement supérieur au traitement standard.

Les définitions « E prometteur » et « E non prometteur », permettent de déterminer des règles d'arrêt de l'étude, en effet l'étude est arrêtée si :

$$\begin{cases} \text{E soit déclaré prometteur} \\ \text{E soit déclaré non prometteur} \\ M_{\max} \text{ atteint} \end{cases}$$

3- Méthodes utilisées

Une solution analytique du critère pour déterminer M est difficilement calculable. En effet, ce critère est une intégrale triple. Une méthode utilisée pour résoudre ce problème est la méthode de Monte Carlo. Ainsi, l'utilisation d'une approximation numérique grâce à la méthode de Monte Carlo permet d'approcher au mieux le résultat de ces intégrales.

Définition 5 : méthode de Monte Carlo

Soit l'intégrale $\int h(\theta) f(\theta) d\theta$, $\theta \in \Theta$, h une fonction connue. Soit $(\theta_1, \dots, \theta_n)$ un échantillon généré suivant la distribution f . L'intégrale précédente est alors approchée par :

$$\frac{1}{n} \sum_{i=1}^n h(\theta_i)$$

Remarque :

Si le paramètre était généré un nombre infini de fois, la valeur exacte de l'intégrale serait obtenue. Ceci est justifié par le théorème de la loi forte des grands nombres. En pratique un grand nombre de tirage suffit pour bien approximer l'intégrale.

Au moment où M est déterminée, aucun patient n'est encore recruté. Il faut alors simuler le nombre de réponses pour M patients permettant de détecter la supériorité du traitement expérimental par rapport au traitement standard.

La méthode des chaînes de Markov est alors utilisée car pour M patients recrutés, M réponses sont possibles (i.e. M états possibles) auxquelles sont associées les distributions a posteriori. Les probabilités d'être dans chacun des états de la chaîne en sont déduites. M est choisie telle que la probabilité d'être dans un état absorbant de la chaîne (traitement expérimental prometteur par rapport au traitement standard) est supérieur à un certain seuil.

Définition 5 : Chaîne de Markov

Soit un espace dénombrable I . Soit $X_0(\omega), X_1(\omega), X_2(\omega), \dots$ une suite de variables aléatoires sur un espace probabilisé (Ω, F, P) à valeurs dans I . Cette suite est une chaîne de Markov, si la loi conditionnelle de la variable aléatoire $X_{n+1}(\omega)$ sachant les valeurs $X_0(\omega) = i_0, \dots, X_n(\omega) = i_n$ ne dépend pas de i_0, \dots, i_{n-1} mais seulement de la dernière valeur i_n .

Soit $p_{i_n, i_{n+1}} = \mathbb{P}(X_{n+1}(\omega) = i_{n+1} | X_0(\omega) = i_0, \dots, X_n(\omega) = i_n)$. Comme la probabilité conditionnelle que $X_{n+1}(\omega) \in I$ est 1, il est clair que $\sum_{i_{n+1}} p_{i_n, i_{n+1}} = 1 \forall i_n$.

4- Calcul de la taille de l'échantillon

Formellement, la méthode pour trouver la taille de l'échantillon est la suivante :

Soit $M_{\min} = 10$ et $M_{\max} = 100$ les tailles minimale et maximale de taille d'échantillon fixé à l'avance.

Pour n patients recrutés, les différents états de la chaîne de Markov sont définis de la manière suivante :

- Les états réponses : $X_n = 0, \dots, X_n = n$
- Les états absorbants : $\begin{cases} p_+ : \text{le traitement expérimental est prometteur} \\ p_- : \text{le traitement expérimental est non prometteur} \end{cases}$

Pour les M_{\min} premiers patients recrutés seules les probabilités d'entrer dans chacun des états réponses sont calculées. En effet, l'essai ne peut pas s'arrêter pour un nombre trop petit de patients i.e. impossibilité de rentrer dans un état absorbant (« E prometteur » ou « E non prometteur »). Ce n'est seulement qu'à partir du $M_{\min}^{\text{ième}}$ patient recruté que les probabilités d'aller dans les états absorbants sont calculées en plus des probabilités des états réponses. La taille de l'échantillon est alors trouvée si la probabilité que le traitement expérimental soit prometteur (p_+) dépasse un certain seuil. (cf. détail de la chaîne de Markov annexe 3)

5- Application

Comme il est décrit précédemment, plusieurs paramètres rentrent en compte dans la détermination de M :

- Le taux de réponse du traitement standard μ_s et l'intervalle autour de ce taux $W_{s,0.9}$
- L'amélioration δ_0 grâce à laquelle le traitement expérimental pourra être déclaré non prometteur par rapport au traitement standard
- Les seuils p_u et p_l permettant de déclarer le traitement expérimental prometteur ou non par rapport au traitement standard
- Le seuil π permettant d'arrêter l'algorithme de simulation et de déterminer M

$\mu_s = 0.2$					
p_u, p_l	$W_{s,0.9}$	δ_0	$\pi = 0.75$	$\pi = 0.80$	$\pi = 0.85$
0.05, 0.95	0.10	0.15	45	69	>100
		0.20	25	31	>100
	0.20	0.15	82	>100	>100
		0.20	32	46	>100
	0.30	0.15	>100	>100	>100
		0.20	71	>100	>100
0.02, 0.98	0.10	0.15	65	81	>100
		0.20	35	42	52

Table 4 : nombre de sujet nécessaire pour $\mu_s = 0.2$

$\mu_s = 0.5$					
p_u, p_l	$W_{s,0.9}$	δ_0	$\pi = 0.75$	$\pi = 0.80$	$\pi = 0.85$
0.05, 0.95	0.10	0.15	49	68	90
		0.20	28	35	30
	0.20	0.15	73	>100	>100
		0.20	32	40	42
	0.30	0.15	>100	>100	>100
		0.20	49	67	>100
0.02, 0.98	0.10	0.15	74	89	>100
		0.20	41	45	55

Table 5 : nombre de sujet nécessaire $\mu_s = 0.5$

Remarquons que plus π est élevé i.e. la puissance est grande, plus le nombre de patients permettant de déclarer le traitement expérimental prometteur sera grand.

Remarquons également que pour un a priori très informatif du traitement standard ($W_{s,0.9} = 0.10$) et pour une puissance de $\pi = 0.85$, les seuils plus précis $(p_u, p_l) = (0.02, 0.98)$ seront préférés.

De manière générale et comme l'indique la figure suivante, il y a toujours un compromis à faire entre la taille de l'échantillon et la puissance que l'on souhaite avoir.

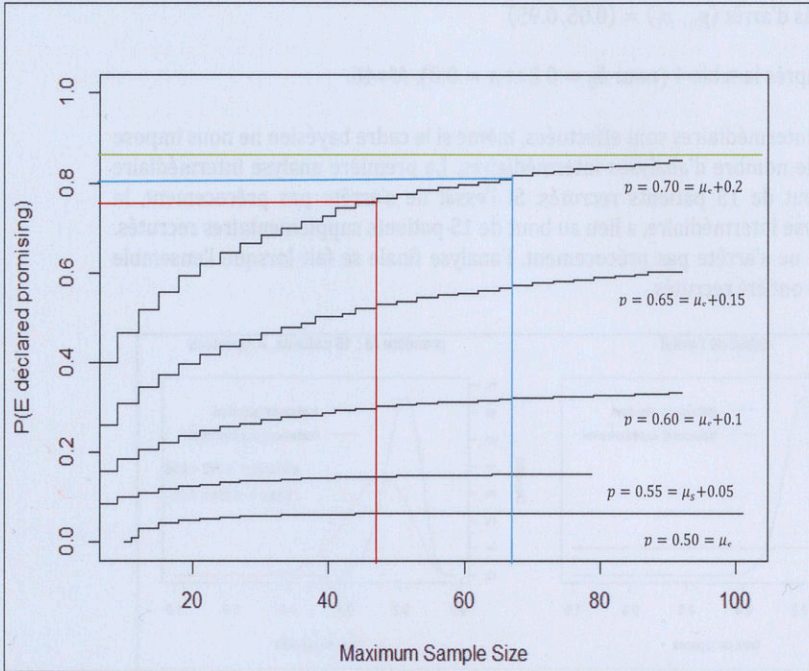


Figure 5 : probabilité de déclarer le traitement expérimental (E) prometteur en fonction de la taille de l'échantillon M , avec la vraie probabilité de succès de E allant de 0.50 à 0.70, $\mu_s = 0.50$, $\delta_0 = 0.20$, $W_{s,90} = 0.30$ et $(p_u, p_l) = (0.05, 0.95)$

La figure précédente illustre comment la probabilité de déclarer le traitement expérimental prometteur évolue en fonction de son vrai taux de réponse et de la taille de l'échantillon. La figure 5 et la table 5 montrent la sensibilité du critère en fonction de M .

3- Les analyses intermédiaires

Au cours de l'étude, lors d'analyses intermédiaires, l'outil bayésien peut également être utilisé pour arrêter l'étude précocement soit pour futilité, soit pour efficacité. Les règles d'arrêt aux analyses intermédiaires sont définies de la façon suivante :

- L'étude s'arrête précocement pour efficacité si : $\mathbb{P}(\theta_s < \theta_E | X_M) > p_u = 0.95$
- L'étude s'arrête précocement pour futilité si : $\mathbb{P}(\theta_s + \delta_0 < \theta_E | X_M) < p_l = 0.05$
- L'étude s'arrête si tous les patients ont été recrutés i.e. si M est atteint.

Pour illustrer l'utilisation de ces règles d'arrêt, supposons que :

- Le taux de réponse du traitement expérimental admet une distribution a priori non informative $\theta_E \sim \beta(1,1)$. Cette distribution sera mise à jour en fonction des informations collectées au moment des analyses intermédiaires.
- Le taux de réponse du traitement standard $\mu_s = 20\%$ avec $W_{s,0.9} = 0.2 \Rightarrow \beta(8,34)$. L'étude est construite avec un seul bras de traitement, ainsi aucune information supplémentaire ne sera collectée au cours de l'essai pour le traitement standard.

- Les seuils d'arrêt $(p_u, p_l) = (0.05, 0.95)$

Dans ce cas, d'après la table 4 (pour $\delta_0 = 0.2$ et $\pi = 0.8$), $M=46$.

Deux analyses intermédiaires sont effectuées, même si le cadre bayésien ne nous impose pas de limite de nombre d'analyses intermédiaires. La première analyse intermédiaire est faite au bout de 15 patients recrutés. Si l'essai ne s'arrête pas précocement, la deuxième analyse intermédiaire, a lieu au bout de 15 patients supplémentaires recrutés. Enfin, si l'essai ne s'arrête pas précocement, l'analyse finale se fait lorsque l'ensemble des 46 patients ont été recrutés.

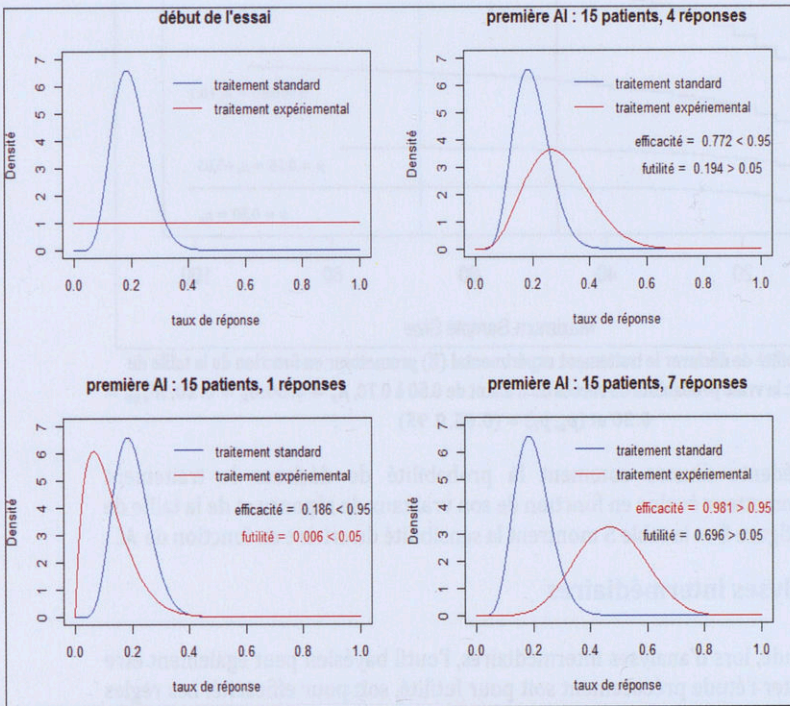


Figure 6 : scénarios possible à la première analyse intermédiaire

Le premier graphique de la figure 6, modélise les connaissances connues à la fois sur le traitement standard et le traitement expérimental au début de l'essai. A la première analyse intermédiaire, la distribution de θ_E est mise à jour en fonction du nombre de réponses x_M pour $M=15$ patients recrutés. L'effet traitement est alors modélisé par la distribution a posteriori pour le traitement expérimental. Plusieurs scénarios sont possibles :

- Peu de réponses sont observées (figure 6: 1 réponse parmi les 15 patients recrutés) et dans ce cas le critère d'arrêt pour futilité est atteint ($\mathbb{P}(\theta_S + \delta_0 < \theta_E | X_M) = 0.012 < p_l = 0.05$)
- Beaucoup de réponses sont observées (figure 6: 7 réponses parmi les 15 patients recrutés) et dans ce cas le critère d'arrêt pour efficacité est atteint ($\mathbb{P}(\theta_S < \theta_E | X_M) = 0.981 > p_u = 0.95$)

- Sinon l'essai ne s'arrête pas précocement à la première analyse intermédiaire. Les 15 patients supplémentaires sont recrutés et nous procédons à la seconde analyse intermédiaire. Comme au moment de la première analyse intermédiaire, plusieurs scénarios sont possibles : arrêt pour futilité, efficacité ou continuation de l'essai (cf. figure 7).

Pour tous les scénarios, remarquons que les distributions a posteriori sont plus précises (plus pointues, resserrées) grâce aux observations disponibles au moment de l'analyse intermédiaire. En effet, l'incertitude a priori sur le traitement diminue, ce qui se traduit par une variation moins grande de la distribution a posteriori.

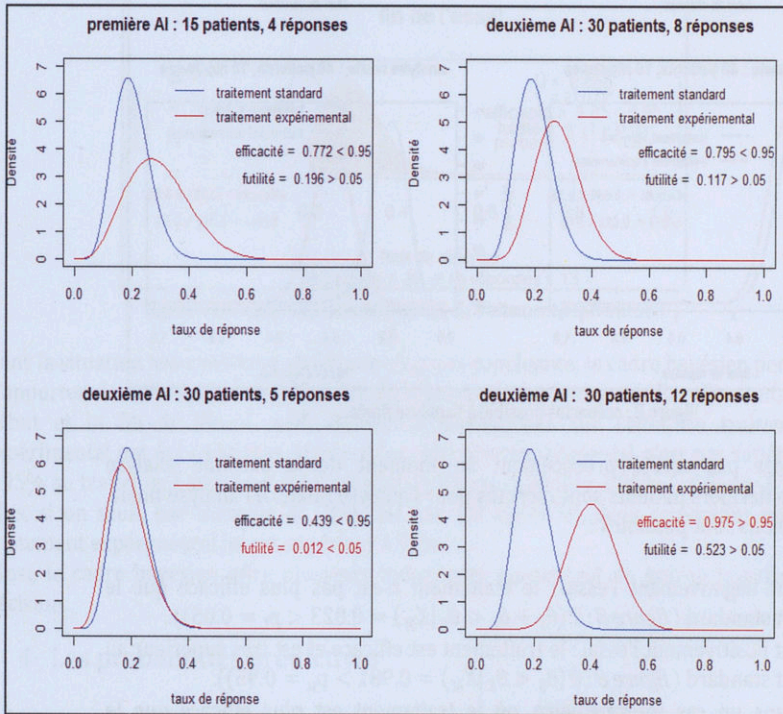


Figure 7 : scénarios possible à la deuxième analyse intermédiaire

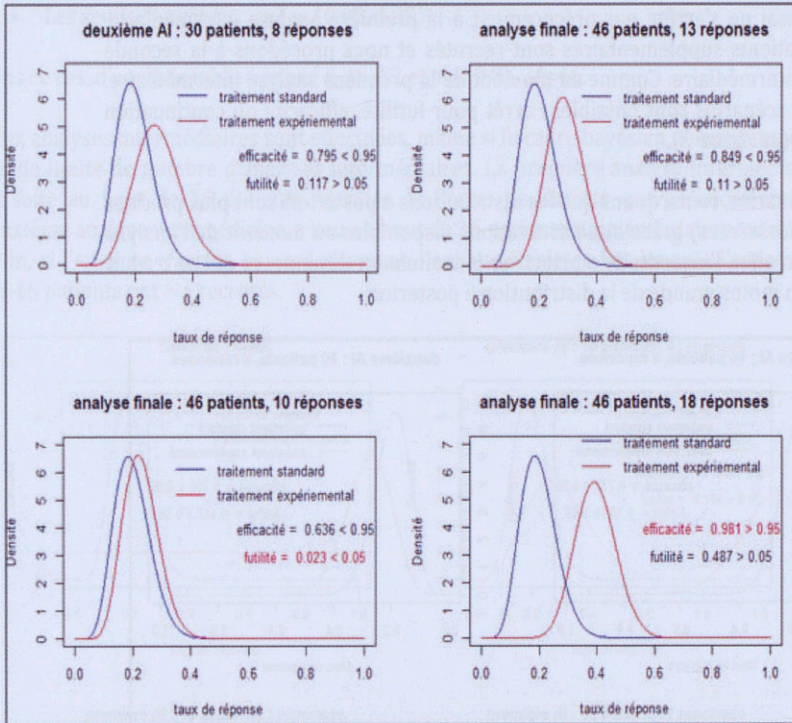


Figure 8 : scénarios possible à l'analyse finale

Si l'essai ne s'est pas arrêté précocement au moment de la seconde analyse intermédiaire, les derniers patients sont recrutés pour l'analyse finale. A l'analyse finale, plusieurs conclusions sont possibles :

- On conclut négativement l'essai : le traitement n'est pas plus efficace que le traitement standard (*figure 8* : $\mathbb{P}(\theta_S + \delta_0 < \theta_E | X_M) = 0.023 < p_l = 0.05$)).
- On conclut positivement l'essai : le traitement est efficace et est très supérieur au traitement standard (*figure 8* : $\mathbb{P}(\theta_S < \theta_E | X_M) = 0.981 > p_u = 0.95$)).
- On est dans un cas intermédiaire, où le traitement est plus efficace que le traitement standard mais le seuil de 95% n'est pas dépassé.

Au travers de cet exemple, l'analyse bayésienne nous donne une réponse claire et facilement interprétable à une question clinique. Contrairement à l'analyse fréquentiste où notre décision serait basée sur une p-value et où deux scénarios seraient possibles (rejet/non rejet de l'hypothèse nulle), ici plusieurs cas de figure sont possibles.

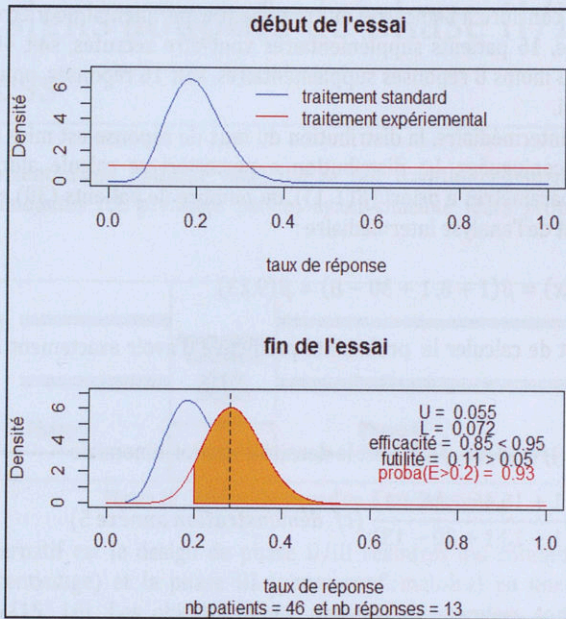


Figure 9 : évolution du taux de réponse du traitement expérimental

Dans la situation intermédiaire où l'étude n'est pas concluante, le cadre bayésien permet d'apporter des réponses sur l'efficacité du traitement expérimental. En effet, entre le début et la fin de l'essai, suffisamment d'informations sur l'effet du traitement expérimental ont été collectées. Même si le traitement expérimental n'est pas supérieur à 95% au traitement standard, il lui est tout de même supérieur à 85% (cf figure 8). De plus, si un seuil, par exemple de 20%, est fixé (cf figure 9), nous constatons que le traitement expérimental lui est supérieur à 93%.

Ainsi, le cadre bayésien offre plusieurs indicateurs permettant de guider la prise de décision.

4- Les probabilités prédictives

Lors d'une analyse intermédiaire ou de l'analyse finale, l'outil bayésien permet de prendre facilement une décision en fonction des observations disponibles. L'analyse bayésienne peut également permettre de faire des prédictions sur les observations futures. En effet, lors d'une analyse intermédiaire ou de l'analyse finale, les probabilités prédictives vont permettre de répondre à des questions telles que :

- Quelle est la probabilité de succès de l'essai sachant les informations disponibles lors de l'analyse intermédiaire ?
- Quelle est la probabilité de succès de la phase III sachant les résultats de la phase II à l'analyse finale ? (critère GO/NO GO)

Reprenons l'exemple du scénario intermédiaire de la figure 7, où lors de la seconde analyse intermédiaire l'essai ne s'arrête pas précocement pour futilité ou efficacité (cf. figure 7 cas : 8 réponses observées pour 30 patients recrutés). Dans le contexte d'un essai clinique planifié, le nombre total de patients à inclure est connu. De plus, le taux de

réponse ciblé pour pouvoir conclure à l'efficacité du traitement expérimental peut être calculé. Pour notre exemple, 16 patients supplémentaires vont être recrutés, soit 46 patients, et il nous faudra au moins 8 réponses supplémentaires, soit 16 réponses, pour conclure positivement l'essai.

Lors de la seconde analyse intermédiaire, la distribution du taux de réponse est mise à jour via les distributions conjuguées. La distribution a posteriori se calcule alors facilement en fonction des paramètres a priori ($\beta(1, 1)$), du nombre de patients (30) et de réponses (13) au moment de l'analyse intermédiaire :

$$\pi(\theta|x) = \beta(1 + 8, 1 + 30 - 8) = \beta(9, 23).$$

La définition 3 nous permet de calculer la probabilité prédictive d'avoir exactement 3 réponses sur 16 patients :

$$\begin{aligned}\mathbb{P}(y|x) &= \int \pi(\theta|x) f(y|\theta) d_{\theta}, f(\cdot|\theta) \text{ est la densité d'une loi binomiale} \\ \mathbb{P}(y|x) &= \binom{16}{8} \frac{B(1 + 16, 46 - 16 + 1)}{B(1 + 13, 1 + 30 - 13)} \text{ (cf. démonstration annexe 5)} \\ \mathbb{P}(y|x) &= 0.05\end{aligned}$$

Nous en déduisons, la probabilité d'avoir au moins 3 réponses sur les 16 futurs patients en fonction de nos connaissances au moment de l'analyse intermédiaire :

$$\begin{aligned}\sum_{i=3}^{16} \mathbb{P}(y \geq i|x) &= \sum_{i=8}^{16} \binom{16}{i} \frac{B(1 + 8 + i, 46 - 13 + i + 1)}{B(1 + 13, 1 + 30 - 13)} \text{ (cf. démonstration annexe 5)} \\ \sum_{i=3}^{16} \mathbb{P}(y \geq i|x) &= 0.092\end{aligned}$$

Le résultat de la probabilité prédictive peut être alors utilisé dans le cadre de design séquentiel ou adaptatif pour prendre la décision ou non de continuer l'essai.

Les designs adaptifs de phase II/III seamless

Le développement d'une nouvelle molécule se fait traditionnellement en plusieurs phases indépendantes (cf. première partie) ayant chacune leurs propres designs et objectifs.

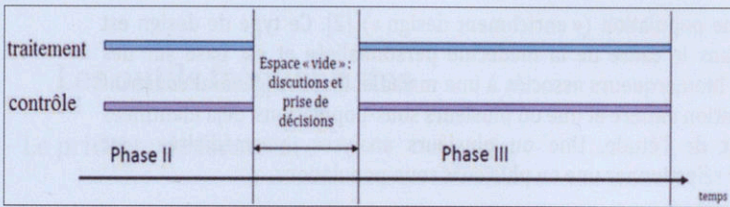


Figure 10 : schéma classique d'un essai clinique

Un design alternatif est le design de phase II/III seamless qui combine la phase IIb (phase d'apprentissage) et la phase III (phase confirmatoire) en une unique phase confirmatoire [15, 16]. Les objectifs d'une phase II/III seamless sont d'établir les premières estimations de l'efficacité thérapeutique du traitement anti-tumoral puis de prouver son efficacité en le comparant à un traitement standard, tout en éliminant le laps de temps qu'il peut y avoir entre les phases II et les phases III. En effet, traditionnellement, il peut se passer plusieurs mois entre ces deux phases, le temps d'analyser les résultats de la phase II et de planifier la phase III.

Les designs adaptatifs de phase II/III seamless [5] sont des designs de phase II/III seamless où les observations des patients recrutés avant et après les adaptations sont utilisées pour l'analyse finale, contrairement aux phases II et III classique où seules les données des patients de la phase III sont utilisées pour établir la preuve formelle de l'efficacité thérapeutique du traitement. Cette prise en compte de tous les patients dans les designs combinés permet possiblement de diminuer le nombre de sujet nécessaire à l'étude ou potentiellement d'atteindre une plus grande puissance qu'un essai traditionnel séparé.

Les adaptations sont faites lors d'analyses intermédiaires, en général au cours de la phase d'apprentissage du design. Plusieurs types d'adaptations peuvent être effectués, notamment :

- Sélectionner un traitement ("pick the winner/drop the loser design") [2]. En général, ces designs sont construits avec plusieurs bras de traitement et une ou plusieurs analyses intermédiaires dont les résultats vont permettre de sélectionner la ou les traitements les plus efficaces.

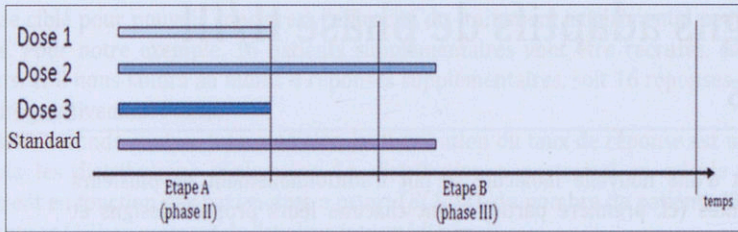


Figure 11 : phase II/III seamless sélection de dose

- Sélectionner une population (« enrichment design ») [2]. Ce type de design est une avancée dans le cadre de la médecine personnalisée et est basé sur des réponses à des biomarqueurs associés à une maladie. Il est en général construit avec une population entière et une ou plusieurs sous-populations, déjà identifiées avant le début de l'étude. Une ou plusieurs analyses intermédiaires sont effectuées pour sélectionner une ou plusieurs sous-populations.

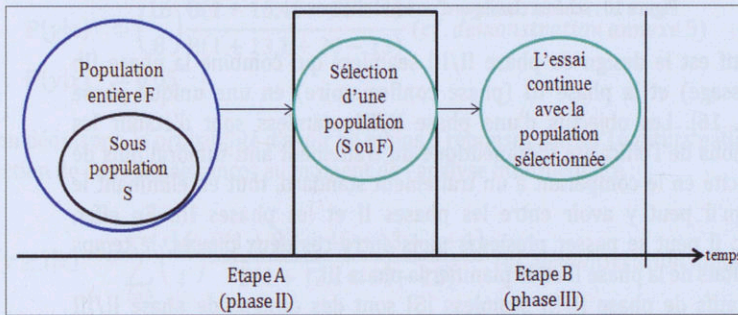


Figure 12 : phase II/III seamless enrichissement de population

- Arrêter l'étude précocement. Les résultats des analyses intermédiaires peuvent permettre d'arrêter l'étude précocement, soit pour futilité si les résultats montrent que tous les traitements sont inefficaces, soit pour efficacité si les résultats montrent qu'au moins un traitement est très efficace.
- Réévaluer la taille de l'échantillon afin d'obtenir une puissance statistique nécessaire pour déterminer l'efficacité thérapeutique du traitement.

Plusieurs méthodologies peuvent être utilisées pour construire ce genre de design :

- Les designs construits grâce à des outils fréquentistes [7]. Ce sont des designs classiques, connus et reconnus par les agences.
- Les designs construits grâce à des outils bayésiens [14]. Ce sont des designs où lors des adaptations aux analyses intermédiaires et lors de l'analyse finale le paradigme bayésien est utilisé. Cependant, les autorités sanitaires et les agences réglementaires exigent le maintien de l'erreur de type I pour les études confirmatoires.
- Les designs construits grâce à des méthodes hybrides bayésiennes/fréquentistes [3, 4, 6, 8, 9]. Ce sont des designs qui utilisent trois méthodologies différentes ; l'approche des tests d'hypothèses adaptatifs, le principe de fermeture (« closure principle ») pour les tests d'hypothèses et contrôle de l'erreur de type I, et les

règles de décision bayésienne basées sur la puissance prédictive et la probabilité a posteriori. Dans ce cas, les méthodes bayésiennes sont uniquement utilisées pour planifier la seconde partie de l'essai.

Dans le cadre du stage, nous nous sommes intéressés aux méthodes hybrides bayésiennes/fréquentistes. Pour pouvoir construire ce genre de design, il nous faut commencer par définir l'ensemble des procédures fréquentistes nécessaires puis définir les outils bayésiens qui nous permettront de prendre des décisions au moment des analyses intermédiaires.

I- Les outils fréquentistes

1- Le principe de fermeture

Le principe de fermeture est un concept important dans le cadre de tests multiples.

Soit n hypothèses nulles $H_i, i = 1, \dots, n$ testées, par exemple n doses comparées à une dose contrôle. Dans le cadre des tests multiples, il est primordiale, pour la validité des tests, de contrôler le taux d'erreur « familywise » (FWER) à un seuil pré-spécifié α . Le FWER contrôle la probabilité de n'avoir aucun faux positif pour un α prédéfini.

	Nb d'hypothèses nulles non rejetées	Nb d'hypothèses nulles rejetées	total
Nb H_0 vraies	U	V : Nb de faux positifs	n_0
Nb H_0 fausses	T : Nb de faux négatifs	S	n_1
total	N-R	R	$N = n_0 + n_1$

Table 6 : erreurs dans les tests d'hypothèses

Ainsi le $\text{FWER} = \mathbb{P}(V > 1)$, c'est-à-dire, c'est la probabilité d'avoir au moins une erreur de type I.

Pour contrôler le FWER, le principe de fermeture considère toutes les intersections construites à partir de l'ensemble des hypothèses nulles initiales. L'hypothèse nulle H_i est alors rejetée au niveau FWER α si toutes les intersections impliquant H_i sont également rejetées. Plus formellement, le principe de fermeture est défini de la façon suivante [8] :

- Soient les hypothèses nulles H_0, \dots, H_n à tester
- Soient toutes les intersections possibles $H_I = \bigcap_{i \in I} H_i, i = 1, \dots, n$
- Pour chaque intersection, un niveau α local est choisi
- H_i est rejetée au niveau FWER α si tous les H_I avec $i \in I$ sont rejetées à un niveau α local

Exemple :

soient trois doses de traitement testées avec une dose contrôle. Nous avons donc 3 hypothèses nulles H_0, H_1 et H_3 . Soient $H_{123} = H_1 \cap H_2 \cap H_3$ et $H_{ij} = H_i \cap H_j, 1 \leq i, j \leq 3$ les 7 intersections possibles définies à partir de l'ensemble initial des hypothèses nulles.

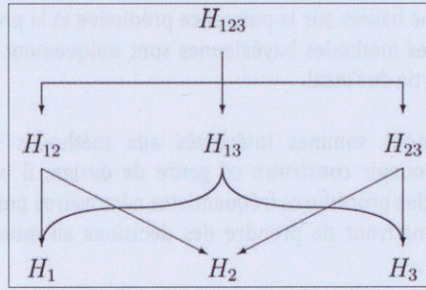


Figure 13 : principe de fermeture pour 3 hypothèses nulles

Dans ce cas, l'hypothèse H_1 , par exemple, est rejetée si H_{123} , H_{12} , H_{13} et H_1 sont rejetées à un niveau local α .

En général, nous commençons par tester l'hypothèse H_{123} et si celle-ci est rejetée, les 2 autres intersections sont testées, sinon, la procédure s'arrête.

Les corrections de Bonferroni, Simes ou Sidak peuvent être utilisées pour calculer le α local.

Définition 6 : Correction de Bonferroni

Soient n hypothèses nulles à tester. Une hypothèse nulle j est rejetée si :

$$p_j \leq \frac{\alpha}{n}, j \in \{1, \dots, n\}$$

où p_j est la p-value non ajustée.

Définition 7 : Correction de Simes

Soient n hypothèses nulles H_0^1, \dots, H_0^n et $H_0 = \bigcap_{i \in n} H_0^i$. Soient $p_1 \leq \dots \leq p_n$ les n p-values ordonnées, non ajustées. H_0 est rejetée dès lors qu'une des n hypothèses nulles est rejetée au niveau $\frac{j \times \alpha}{n}$.

Définition 8 : Correction de Sidak

Soient n hypothèses nulles à tester. Une hypothèse nulle j est rejetée si :

$$p_j \leq 1 - (1 - \alpha)^{1/n}, j \in \{1, \dots, n\}$$

où p_j est la p-value non ajustée.

2- La procédure de test adaptatif

Pour plus de simplicité, plaçons-nous dans le cadre d'un essai clinique en deux étapes comparant un traitement expérimental à un traitement contrôle et où une analyse intermédiaire est réalisée. Les données collectées lors de l'analyse intermédiaire sont utilisées pour prendre les décisions de continuer ou d'arrêter précocement l'étude. Si l'essai ne s'arrête pas précocement, l'analyse finale utilise les résultats des deux étapes. Soient $p_i, i \in \{1, 2\}$ les p-values issues des tests effectués à la fin des deux étapes, la procédure de test adaptatif [8] est la suivante :

- La procédure de test pour l'étape 1 et les règles d'arrêt sont définies, la fonction de combinaison des p-values pour l'analyse finale est pré-spécifiée
- A la fin de la première étape la p-value p_1 du test est calculée. Grâce à cette valeur, nous décidons de continuer ou d'arrêter l'essai.
- Si l'essai ne s'arrête pas, la p-value p_2 du test de la deuxième étape est calculée (indépendant de p_1 sous l'hypothèse nulle i.e. vérifie la condition *p-clud*)
- Calcul de la fonction de combinaison à l'analyse finale.

Plusieurs tests de combinaison peuvent être utilisés :

- Le test de combinaison de Fisher : $C(p_1, p_2) = p_1 \cdot p_2 \leq c = e^{-\chi_{4,1-\frac{\alpha}{2}}^2}$
- Méthode de la normale inversée :

$$C(p_1, p_2) = 1 - \Phi(w_1 \Phi^{-1}(1 - p_1) + w_2 \Phi^{-1}(1 - p_2)), \text{ où } 0 < w_i < 1 \\ \sum w_i^2 = 1 \text{ et } \Phi \text{ est la fonction de répartition d'une } \mathcal{N}(0,1)$$

- Fonctions d'erreur conditionnelle :

$$\mathbb{P}(\text{rejeter } H_0 | H_1) = \begin{cases} 1 & \text{si } p_1 \leq \alpha_1 \\ 0 & \text{si } p_1 \geq \alpha_0 \\ \max\{p_2 | C(p_1, p_2) \leq c\} & \text{si } p_1 \in]\alpha_1, \alpha_0[\end{cases}$$

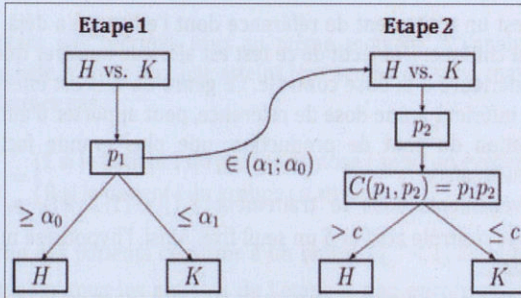


Figure 14 : design adaptatif en deux étapes

La combinaison des deux procédures fréquentistes détaillés précédemment permet de construire les designs adaptatifs avec plusieurs hypothèses nulles, c'est la « procédure de Hommel » ou encore le « test de combinaison adaptatif » [8]. Ce sont donc des méthodologies utilisées pour les designs adaptatifs de phase II/III seamless entièrement fréquentistes, où les adaptations seront basées sur les résultats des tests statistiques. Cette méthodologie est également utilisée pour le contrôle de l'erreur de type I et pour l'analyse finale dans le cadre de design hybride bayésien/fréquentiste. Les adaptations telles que la sélection de sous-population ou de traitement ainsi que l'arrêt de l'étude seront basés sur le paradigme bayésien. Pour les designs hybrides, les puissances prédictives, permettant de prendre des décisions, sont calculées en fonction des tests statistiques.

II- Calcul de la puissance prédictive

1- Le contexte

Nous allons calculer les puissances prédictives dans le cadre d'un design adaptatif de phase II/III seamless avec une analyse intermédiaire. C'est lors de cette analyse intermédiaire que la puissance prédictive sera calculée. Elle permettra, dans ce contexte, de sélectionner une dose de traitement. La méthode utilisée provient de l'article [4].

On suppose avoir k doses de traitement comparées à une dose contrôle. L'étude se déroule en deux étapes. Une étape A durant laquelle les k doses et la dose contrôle seront observées et à la fin de laquelle la dose la plus prometteuse sera sélectionnée; et une étape B durant laquelle la dose sélectionnée et la dose contrôle seront observées et à la fin de laquelle nous établiront, en utilisant la procédure de test adaptative, la preuve formelle de l'efficacité ou non de la dose sélectionnée.

Ici, le contexte sera celui des données de survie où pour chaque dose, le nombre d'événements sera comparé au nombre d'événement de la dose contrôle.

Les tests statistiques utilisées pour comparer les k doses à la dose contrôle dans un premier temps, puis la dose sélectionnée à la dose contrôle dans un second temps, sont les tests de non-infériorité et de supériorité.

a- Test de non-infériorité

Dans ce cas, la dose contrôle est un traitement de référence dont l'efficacité a déjà été prouvé dans le cadre d'un essai clinique. L'objectif de ce test est alors de montrer que la nouvelle dose n'est pas non inférieure à la dose contrôle. Ce genre de test est effectué lorsqu'une nouvelle dose, non inférieure à une dose de référence, peut apporter d'autres avantages tels que la diminution du coût de production, une plus grande facilité d'administration ou une meilleure safety.

Soient μ_j la proportion d'événements sous le traitement j ($j \in \{1, \dots, k\}$), μ_0 la proportion d'événements sous le contrôle actif et δ un seuil fixé. Ainsi, l'hypothèse nulle sera définie de la manière suivante :

$$H_{Nj}: \mu_j - \mu_0 \geq \delta, \delta > 0$$

b- Test de supériorité

Dans ce cas, nous voulons démontrer que la nouvelle dose a un effet traitement significativement supérieur à la dose contrôle. Ainsi, l'hypothèse nulle sera définie de la manière suivante :

$$H_{Sj}: \mu_j - \mu_0 \geq 0$$

c- Le jeu de donnée

A la fin de l'essai, le jeu de donnée entier disponible avec lequel nous effectueront l'ensemble des tests est le suivant :

$$\begin{cases} (t_{Aij}, y_{Aij}) & \text{pour le patient } i \text{ du groupe de dose } j \text{ de l'étape A} \\ (t_{Bij}, y_{Bij}) & \text{pour le patient } i \text{ du groupe de dose } j \text{ de l'étape B} \end{cases}$$

où t_{Kij} , (avec $K = A, B$), est le temps d'évènement ou de censure et :

$$y_{Kij} = \begin{cases} 1 & \text{si le patient } i \text{ du groupe de dose } j \text{ subit un évènement} \\ 0 & \text{sinon} \end{cases}$$

Si le patient ne subit pas d'évènement lors de son temps de suivi alors $t_{Kij} = t_0$, où t_0 est le temps de suivi maximal d'un patient. Un patient est suivi jusqu'à ce qu'il subisse un évènement ou qu'il atteigne t_0 .

Les patients recrutés à l'étape A sont assignés à $j = 0, \dots, k$ groupe de dose, où la dose 0 est la dose contrôle ; tandis que les patients recrutés à l'étape B sont assignés à $j = 0$ ou $j = j_0$ la dose sélectionnée lors de l'analyse intermédiaire.

Lors de l'analyse intermédiaire (à la fin de l'étape A), seules les observations de l'étape A sont disponibles et peuvent être partielles. Dans ce cas, nous définissons $z = (z_{ij}) = (t_{Aij}', c_{Aij})$ le jeu de données disponible à l'analyse intermédiaire, où $t_{Aij}' \leq t_{Aij}$ et :

$$c_{Aij} = \begin{cases} 1 & \text{si le patient } i \text{ du groupe de dose } j \text{ est censuré à l'analyse intermédiaire} \\ 0 & \text{sinon} \end{cases}$$

Les patients n'étant pas recrutés tous au même moment, à l'analyse intermédiaire seulement une partie d'entre eux ont atteint leur temps de suivi maximal ou ont subi l'évènement et dans ce cas :

$$y_{Aij} = \begin{cases} 1 & \text{si le patient } i \text{ du groupe de dose } j \text{ subit un évènement} \\ 0 & \text{si le patient } i \text{ du groupe } j \text{ a atteint } t_0 \end{cases}$$

Pour l'autre partie des patients censurés à un temps $t_{Aij}' < t_0$ au moment de l'analyse intermédiaire et pour tous les patients de l'étape B non encore recrutés la valeur de y_{Kij} est manquante.

La figure suivante décrit le suivi des patients pour deux doses et la dose contrôle à l'étape A, la dose sélectionnée et la dose contrôle à l'étape B. Les patients sont suivis jusqu'à subir l'évènement (x) ou atteindre le temps de suivi maximal. Au moment de l'analyse intermédiaire, tous les patients de l'étape A sont recrutés mais une partie d'entre eux n'ont ni subi l'évènement, ni atteint le temps maximal de suivi.

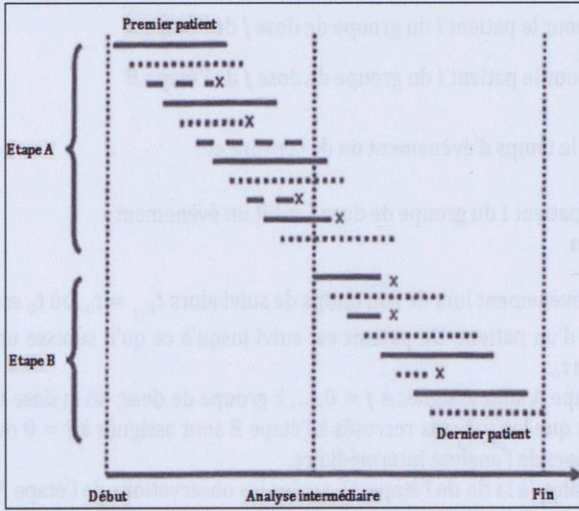


Figure 15 : schéma de suivi des patients

2- La puissance prédictive

C'est lors de l'analyse intermédiaire que la puissance prédictive est calculée. Elle permet de prendre des décisions, sans influencer sur l'erreur de type I, telles que :

- sélectionner la dose la plus prometteuse parmi les k doses de traitement testées
- arrêter l'étude pour futilité si aucune des doses testées ne semblent prometteuses par rapport à la dose contrôle
- réévaluer le nombre de patients pour l'étape B de l'étude

La puissance prédictive est un concept hybride utilisant l'approche bayésienne tout en gardant la notion de test de fréquentiste pour l'analyse des données :

$$PP = \int CP(\theta)\pi(\theta|\text{données à l'analyse intermédiaire})d_{\theta}$$

Où $CP(\theta)$ est la puissance conditionnelle définie comme étant la probabilité d'avoir un résultat statistiquement significatif à la fin de l'étude sachant les données disponibles lors de l'analyse intermédiaire.

Le paramètre d'intérêt θ n'est pas fixe et inconnu comme en fréquentiste, mais admet une distribution qui témoigne à la fois de notre connaissance et de notre incertitude vis-à-vis de ce dernier.

Pour pouvoir calculer la puissance prédictive, un modèle paramétrique doit être défini.

a- Le modèle paramétrique

Pour rendre le calcul de la puissance prédictive opérationnelle, nous définissons un modèle paramétrique pour notre paramètre d'intérêt. La méthode hybride bayésienne/fréquentiste a été, au travers de ce stage, étudiée dans le cadre des données de survie. Nous utilisons alors le modèle exponentiel qui est le principal modèle

paramétrique en analyse de données de survie. Ainsi, les temps de survie t_i sont indépendants et identiquement distribués selon une loi exponentielle de paramètre θ . Dans un cadre général :

$$t_i \sim \mathcal{E}(\theta) \text{ avec } f(t_i, \theta) = \theta e^{-\theta t_i} \text{ la densité de } t_i$$

Soit la fonction de survie :

$$S(t_i, \theta) = e^{-\theta t_i} \text{ avec } \theta \sim \mathcal{G}(\alpha_0, \beta_0) \text{ la distribution a priori du taux de hasard}$$

Lors de l'analyse intermédiaire, grâce aux observations disponibles, la distribution a priori de notre paramètre est mise à jour [1] :

$$\theta | \text{obs} \sim \mathcal{G}(\sum_{i=0}^n c_i + \alpha_0, \sum_{i=0}^n t_i + \beta_0), \text{ où } c_i \text{ est l'indicateur de censure (cf. démonstration annexe 6)}$$

Dans un cadre moins général, utilisé lors de ce stage, le temps de suivi de chaque patient est découpé en morceaux :

$$s_0 = 0 < s_1 < s_2 < \dots < s_m = t_0$$

Le modèle de hasard est alors constant par morceau (« piecewise constant hazard model ») ou encore appelé distribution exponentielle par morceaux (« piecewise exponential distribution »). Ce modèle est flexible et facilement interprétable pour la fonction de survie. La probabilité de subir un événement au temps $t \in]s_{r-1}, s_r]$, $r = 1, \dots, m$, pour une certaine dose de traitement est définie de la manière suivante :

$$S(t, \theta) = e^{-\theta_1 s_1} e^{-\theta_2 (s_2 - s_1)} \dots e^{-\theta_r (t - s_{r-1})}$$

où $\theta_i \sim \mathcal{G}(\alpha_i, \beta_i)$ défini pour chaque intervalle de temps

La mise à jour de la distribution du paramètre en fonction des observations disponibles au moment de l'analyse intermédiaire, s'effectue de la façon suivante:

Si le patient a subi un événement ($y = 1$) ou si le patient est censuré ($y = 0$) au temps $t \in]s_{r-1}, s_r]$ alors :

$$\begin{cases} \theta_r | \text{obs} \sim \mathcal{G}(a_r + y, b_r + (t - s_{r-1})) \\ \theta_u | \text{obs} \sim \mathcal{G}(a_u + y, b_u + (s_u - s_{u-1})) \text{ si } u < r \\ \theta_u | \text{obs} \sim \mathcal{G}(a_u, b_u) \text{ si } u > r \end{cases}$$

En répétant cette procédure pour chaque patient, la distribution a posteriori du paramètre est obtenue.

Pour les designs adaptatifs seamless, la distribution a priori des taux de hasard par morceaux est définie pour chaque groupe de dose. A l'analyse intermédiaire les distributions sont mises à jour analytiquement en fonction des observations. Basée sur ces distributions a posteriori, nous calculons la puissance prédictive pour les différentes adaptations considérées.

b- Calcul de la puissance conditionnelle

Soit les paramètres a posteriori $\theta = (\theta_j), j = 1, \dots, m$ de la fonction de survie. Soit la puissance conditionnelle :

$$CP(\theta, j_0) = \int R(y_{obs}, y_{mis}) \pi(y_{mis} | \theta, z, j_0) d_{y_{mis}}$$

Où :

$$R(y_{obs}, y_{mis}) = \begin{cases} 1 & \text{si le résultat basé sur } (y_{obs}, y_{mis}) \text{ est statistiquement significatif} \\ 0 & \text{sinon} \end{cases}$$

Et $\pi(y_{mis} | \theta, z, j_0)$ est la densité des observations manquantes lors de l'analyse intermédiaire, définie en fonction θ et de z les observations disponibles pour la dose j_0 . Cette densité est calculée comme il suit :

$$\pi(y_{mis,ij} | \theta, z, j_0) \sim \begin{cases} B(S(t_{A_{ij}}'; \theta_j) - S(t_0; \theta_j)) & \text{pour le patient } i \text{ du groupe } j \text{ de l'étape A} \\ & \text{censuré à l'AI et avec un temps de censure } t_{A_{ij}}' < t_0 \\ B(1 - S(t_0; \theta_j)) & \text{pour le patient } i \text{ du groupe } j \text{ de l'étape B} \end{cases}$$

c- Calcul de la puissance prédictive

La puissance prédictive est définie par :

$$PP(j_0) = \int CP(\theta, j_0) \pi(\theta | z) d_{\theta}$$

Une solution analytique pour ce genre d'intégrale est difficile à obtenir, la méthode de Monte Carlo est alors utilisée pour approximer l'intégrale :

Soit l'échantillon $\theta(1), \dots, \theta(M)$ tiré à partir de la distribution a posteriori $\pi(\theta | z)$ pour un bras de dose. La puissance prédictive est obtenue grâce :

$$PP(j_0) \approx \frac{1}{M} \sum_{u=1}^M CP(\theta(u), j_0)$$

Une solution analytique pour la puissance conditionnelle est également difficilement calculable. Soit plusieurs jeux de données manquants $y_{mis, \theta(u)}(1), \dots, y_{mis, \theta(u)}(N)$ générés suivant la distribution $\pi(y_{mis} | \theta(u), z, j_0)$. La puissance conditionnelle est approchée par :

$$CP(\theta(u), j_0) \approx \frac{1}{N} \sum_{v=1}^N R(y_{obs}, y_{mis, \theta(u)}(v))$$

Donc :

$$PP(j_0) = \int CP(\theta, j_0) \pi(\theta|z) d\theta$$

$$PP(j_0) = \int \left(\int R(y_{obs}, y_{mis}) \pi(y_{mis} | \theta, z, j_0) d_{y_{mis}} \right) \pi(\theta|z) d\theta$$

$$PP(j_0) = \iint R(y_{obs}, y_{mis}) \pi(y_{mis} | \theta, z, j_0) \pi(\theta|z) d_{y_{mis}} d\theta$$

$$PP(j_0) \approx \frac{1}{M} \sum_{u=1}^M CP(\theta(u), j_0)$$

$$PP(j_0) \approx \frac{1}{M} \frac{1}{N} \sum_{u=1}^M \sum_{v=1}^N R(y_{obs}, y_{mis, \theta(u)}(v))$$

La puissance prédictive ainsi calculée va être un outil d'aide à la décision.

Exemples

I- Application

1- Le contexte

Dans le cadre du stage, nous avons repris et implémenté grâce au logiciel R l'exemple de l'article [4]. Cet article se place dans le contexte de développement d'un régime immunosuppresseur dans la transplantation du rein de novo. Le critère principal de jugement est le rejet, la perte de greffe ou la mort au bout de 6 mois après la transplantation. Pour éviter ces événements, un régime courant est la triple immunosuppression avec la cyclosporine, les corticostéroïdes et l'acide mycophénolique. L'efficacité de cette thérapie a été prouvée lors d'essais de phase III classiques comparée avec un comparateur actif. Les résultats de ces essais montrent un taux d'événement à 6 mois de 25 à 30 %.

Une nouvelle thérapie immunosuppressive est développée. Le but est de montrer la non infériorité et possiblement la supériorité de cette nouvelle thérapie par rapport à la triple immunosuppression cyclosporine, corticostéroïdes et acide mycophénolique. Trois doses de cette nouvelle thérapie sont considérées

Pour un essai traditionnel séparé de phase II puis de phase III, 80 patients puis 230 patients par bras de doses sont recrutés respectivement. Le nombre de patients pour la phase II est choisi afin d'avoir une estimation suffisamment précise du taux d'événement à 6 mois. Les patients de la phase II sont recrutés à raison de 40 patients par mois. La phase III a lieu 6 mois après la fin de la phase II et est construite avec la dose sélectionnée et la dose contrôle. Cette étude ainsi construite aurait duré 38 mois et aurait nécessité le recrutement de 780 patients.

Pour un design adaptatif de phase II/III Seamless, lors de l'étape A, 120 patients par bras de dose sont recrutés. L'analyse intermédiaire a lieu 1 an après le recrutement du premier patient, à ce moment de l'étude l'ensemble des patients prévu pour la l'étape A, c'est-à-dire 480 patients, ont été recrutés et seulement la moitié d'entre eux ont été observés pendant 6 mois. L'étape B suit directement l'étape A de l'essai et 120 patients par bras sont recrutés (la dose sélectionnée et la dose contrôle). Ainsi, pour cette étude un total de 720 patients et une durée de 2 ans sont nécessaires.

2- Le calcul de la puissance prédictive

La puissance prédictive se calcule au moment de l'analyse intermédiaire pour concevoir au mieux l'étape B de l'étude.

Supposons qu'au bout de 6 mois de suivi sans événement, la transplantation est un succès. 6 mois est alors défini comme étant le temps de suivi maximal pour chaque patient et chaque bras de dose. Ce temps de suivi est découpé en 4 morceaux :

$$s_0 = 0 < s_1 = 0.5 < s_2 = 1 < s_3 = 3 < s_4 = t_0 = 6 \text{ mois}$$

Soit la fonction de survie à 6 mois pour un bras de dose:

$$S(t_0, \theta) = \exp \left(- \sum_{i=1}^4 \theta_i (s_i - s_{i-1}) \right)$$

Les paramètres $\theta_i, i = 1, \dots, 4$, admettent une distribution pour chaque bras de dose.

a- Distribution a priori

Une première étape, avant le début de l'essai, est d'élucider les paramètres de la distribution a priori des taux de hasard pour chaque bras de dose. Pour chacune des doses les taux de hasard suivent une distribution gamma :

$$\theta_i^{\text{contrôle}} \sim \mathcal{G}(a_i, b_i) \text{ et } \theta_i^{\text{dose } j} \sim \mathcal{G}(\alpha_i, \beta_i)$$

Pour la thérapie immunosuppressive classique, i.e. la dose contrôle, les essais antérieurs apportent suffisamment d'informations pour avoir une distribution précise des taux de hasard sur les 4 intervalles de temps. Cette thérapie contrôle a un taux d'événement de 30% à 6 mois avec une largeur de 0.15 centrée autour de ce taux. La moitié des événements ont lieu lors du premier intervalle de temps, soit 15% des événements, et l'autre moitié des événements est réparti de manière équitable dans les 3 autres intervalles de temps.

Pour la distribution des taux de hasard des trois nouvelles doses, seules les essais de phase préclinique, de phase I et de phase IIa apportent de l'information pour modéliser ces distributions. Ces trois doses ont un taux d'événement à 6 mois de 30% mais avec un intervalle de crédibilité moins précis de largeur 0.30 autour de ce taux.

Les paramètres des distributions gamma des doses expérimentales sont choisis tels que :

- La moyenne a priori pour les doses expérimentales et la dose contrôle est la même pour chaque intervalle de temps

$$\mathbb{E}(\theta_i^{\text{contrôle}}) = \frac{a_i}{b_i} = \mathbb{E}(\theta_i^{\text{dose } j}) = \frac{\alpha_i}{\beta_i}$$

- La probabilité d'avoir un taux d'événement cumulé à 6 mois de 15% est de 5%
- le même coefficient de variation (α_i) est utilisé pour chaque intervalle de temps

Intervalles de temps	Traitement contrôle			Traitement expérimental		
	a	b	F (moyenne, quantile 5% et 95%)	a	b	F (moyenne, quantile 5% et 95%)
]0, 0.5]	9	28	15%(8%,23%)	2	6	15%(3%,32%)
]0.5, 1]	9	74	20%(13%,28%)	2	16	20%(7%,37%)
]1, 3]	9	279	25%(18%,33%)	2	62	25%(11%,42%)
]3, 6]	9	391	30%(23%,38%)	2	87	30%(15%,45%)

Table 7 : paramètres a priori des taux de hasard pour chaque intervalle de temps et la fonction de répartition

b- Jeux de données à l'analyse intermédiaire

N'ayant pas accès à un jeu de données réelles, les observations disponibles lors de l'analyse intermédiaire ont été simulées. Pour simuler ce jeu de données, pour chaque bras de dose, les informations suivantes sont utilisées :

- 4 bras de traitement de 120 patients chacun à l'étape A
- Au moment de l'analyse intermédiaire, la moitié des patients ont été censuré à 6 mois, l'autre moitié étant censuré à un temps strictement inférieur à 6 mois ou ayant subi l'événement. Pour la seconde moitié des patients :

$$\begin{cases} \text{si } y_{Aij} = 1 \text{ le temps est généré suivant une loi uniforme en respectant les taux} \\ \quad \text{d'événements pour chaque intervalle de temps (table 7)} \\ \text{si } y_{Aij} = 0 \text{ le temps est strictement inférieur à 6 mois} \end{cases}$$

Pour la première moitié des patients $y_{Aij} = 0$ avec un temps de suivi égal à 6 mois.

Arbitrairement un nombre d'événement a été choisi pour chaque groupe de traitement afin d'avoir des scénarios différents :

Traitement	Nombre d'événements
Contrôle (dose 1)	$\sum_{i=1}^{120} y_{Ai1} = 36$
Dose 2	$\sum_{i=1}^{120} y_{Ai2} = 30$
Dose 3	$\sum_{i=1}^{120} y_{Ai3} = 42$
Dose 4	$\sum_{i=1}^{120} y_{Ai4} = 48$

Table 8 : nombre d'événement dans chaque groupe de traitement à l'analyse intermédiaire

Ce jeu de données ainsi simulé est considéré comme étant les observations disponibles au moment de l'analyse intermédiaire, à la fin de l'étape A. Il va permettre dans un premier de mettre à jour les distributions des taux de hasard afin de simuler les observations manquantes à l'analyse intermédiaire puis de calculer la puissance prédictive.

c- Simulation des données manquantes

Soient les paramètres des taux de hasard mis à jour dans chaque intervalle de temps et pour chaque dose en fonction du jeu de données comme développé dans la section 2.a du chapitre précédent:

traitement	Distribution des taux de hasard a posteriori		
	Intervalle de temps	a	b
Contrôle (dose 1)]0,0.5]	27	85
]0.5,1]	15	123
]1,3]	15	451
]3,6]	15	597
Dose 2]0,0.5]	17	63
]0.5,1]	7	66
]1,3]	7	238
]3,6]	7	299
Dose 3]0,0.5]	23	62
]0.5,1]	9	63
]1,3]	9	229
]3,6]	9	290
Dose 4]0,0.5]	26	61
]0.5,1]	10	60
]1,3]	10	217
]3,6]	10	291

Table 9 : distribution a posteriori des taux de hasard

Pour chaque bras de dose, les taux de hasard sont mis à jour en fonction du nombre d'événements et de censures dans chaque intervalle de temps. Comme vu précédemment, la distribution des taux de hasard à posteriori est une loi conjuguée. Elle s'écrit en sommant le nombre de censures et d'événements avec le paramètre α_i a priori, et en sommant les temps de censures et d'événements avec le paramètre a priori β_i . Nous en déduisons alors les fonctions de survie a posteriori pour chaque bras de dose. Grâce à ces fonctions de survie, les observations manquantes pour une partie des patients recrutés de l'étape A censurés à moins de 6 mois strictement et l'ensemble des patients non encore recrutés de l'étape B sont simulées.

Une fois les données manquantes simulées, un jeu de données complet est disponible. Nous sommes alors dans les conditions de fin d'étude où tous les patients ont été recrutés et ont soit subi l'événement soit censurés à 6 mois.

Les tests de non infériorité et supériorité peuvent maintenant être effectués.

d- Résultats

La méthode utilisée pour calculer la puissance prédictive est celle détaillée dans la section 2.c du chapitre précédent. Cette méthode est basée sur des simulations du jeu de données manquant en fonction des observations disponibles. En effet une solution analytique pour la puissance prédictive est difficilement calculable. Nous devons alors simuler un certain nombre de jeux de données pour chaque taux de hasard tiré aléatoirement et ce pour chaque bras de doses :

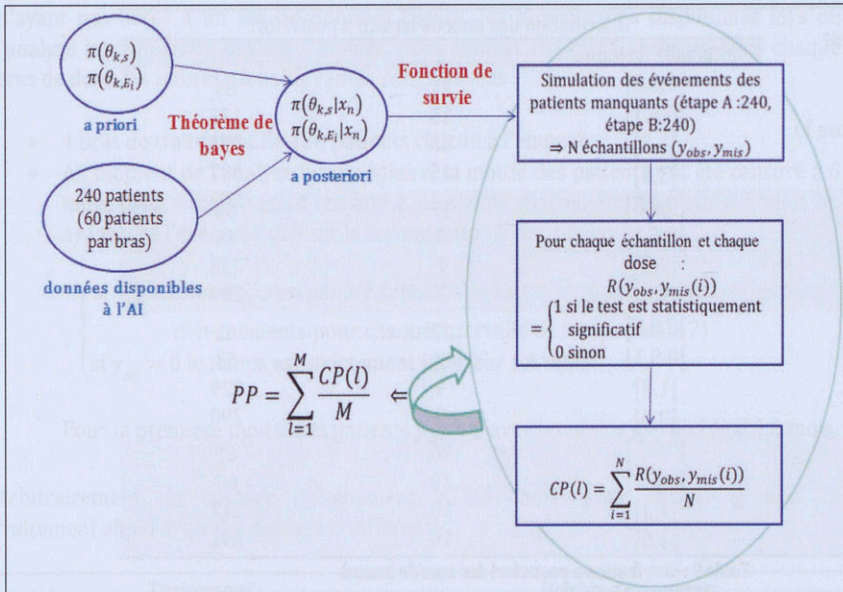


Figure 16 : calcul de la puissance prédictive

Pour calculer la puissance prédictive (PP), nous avons tiré aléatoirement $M=50$ θ_k , $k \in \{1,2,3,4\}$ différents pour chaque bras de dose et pour lesquelles nous avons simulé $N=500$ jeux de données « manquants » afin d'approcher l'intégrale de la PP. Pour le calcul des tests de non-infériorité et de supériorité, nous avons utilisé la correction de Bonferroni dans la procédure de test adaptatif et un $\delta = 0.12$ pour le test de non-infériorité.

traitement	F (moyenne, quantile) a priori	F (moyenne, quantile) a posteriori	PP test de NI	PP test de SUP
Contrôle (dose 1)	30%(23%,38%)	30%(25%,35%)		
Dose 2	30%(16%,45%)	27%(21%,33%)	0.96(0.80,1.00)	0.10(0.00,0.41)
Dose 3	30%(16%,45%)	35%(28%,41%)	0.21(0.01,0.73)	0.01(0.00,0.04)
Dose 4	30%(16%,45%)	39%(32%,45%)	0.09(0.00,0.34)	0.00(0.00,0.01)

Table 10 : taux d'événement à 6 mois a posteriori lors de l'analyse intermédiaire, puissance prédictive pour les tests de non infériorité et de supériorité

La PP est la probabilité d'avoir un test statistiquement significatif à la fin de l'essai pour les tests de non infériorité ou de supériorité. Pour cet exemple, la dose 2 a une PP de 96% pour le test de non-infériorité et donc serait la dose à sélectionner pour aller à l'étape B. Cette PP est beaucoup plus grande que pour les deux autres doses. Ainsi, la dose 2 peut être sélectionnée avec une grande confiance dans le fait que le l'essai sera un succès pour la non infériorité. Si la dose 2 est sélectionnée, nous aurons 96% de chance, au vu de ce qui a été observé, de conclure positivement à la fin de l'étude pour la non infériorité. Cependant, si la dose 2 est sélectionnée, il y a peu de chance de conclure à la supériorité de cette dose (10%).

Si pour des raisons de sécurité, les doses 2 et 3 ne sont pas sélectionnées, l'étude s'arrêterait pour futilité. En effet, la PP pour la dose 4 est suffisamment faible pour le test de non-infériorité et de supériorité, pour dire avec suffisamment de confiance que l'étude ne sera pas un succès si cette dose est sélectionnée.

Plusieurs scénarios peuvent également être comparés. Nous utilisons les mêmes distributions a priori que précédemment. Supposons que :

- La dose avec la plus grande PP est sélectionnée pour l'étape B de l'étude
- On arrête l'essai pour futilité si toutes les PP sont strictement inférieure à 20%

Pour cet exemple, pour chaque scénario, nous avons tiré aléatoirement 25 θ_i , $i \in \{1,2,3,4\}$ différents pour chaque bras de dose et pour lesquelles 200 jeux de données « manquants » ont été simulés afin d'approcher l'intégrale de la PP. Contrairement à l'exemple précédent, où ce processus a été fait une seule fois, ici nous réitérons le calcul de la PP 100 fois afin de voir la stabilité ce genre de design.

Comme précédemment, pour le calcul des tests de non-infériorité et de supériorité, nous avons utilisé la correction de Bonferroni dans la procédure de test adaptatif et un $\delta = 0.12$ pour le test de non-infériorité.

Taux d'événement à 6 mois a posteriori (dose contrôle : 31%)	Dose sélectionnée(%)			Essais sans arrêt pour futilité	Moyenne de la PP (NI) pour la dose sélectionnée
	1	2	3		
40 40 40	37	33	30	30	15
35 35 35	35	33	32	90	40
31 31 31	35	35	30	100	80
23 31 38	97	3	0	100	98
27 31 35	80	20	0	100	85
30 32 35	70	25	5	100	80

Table 11 : caractéristiques d'un design adaptatif de phase II/III seamless utilisant la PP pour sélectionner la meilleure dose ou arrêter l'essai pour futilité

Les résultats suggèrent que la procédure de décision est cohérente. En effet, plus la moyenne de la PP pour le test de non-infériorité est grande plus la fréquence d'essai arrêté pour futilité est petite.

II- Perspectives : l'enrichissement de population

L'émergence de nouvelles thérapies, comme les thérapies ciblées et les immunothérapies en oncologie confrontent les industriels de la santé à développer des designs différents des designs classiques. En effet, sélectionner la mauvaises sous-population peut compromettre le succès de l'étude et limiter l'accès du traitement à une partie de la population pour laquelle il serait bénéfique, tandis que sélectionner la « bonne » sous-population avec un degré acceptable de confiance peut permettre non

seulement d'augmenter les chances de succès de l'étude mais aussi d'en diminuer le temps et les coûts. De plus, ignorer une sous-population peut affaiblir l'effet traitement qui sera dilué dans la population entière.

Dans ce contexte, les designs adaptatifs de phase II/III seamless utilisant des outils bayésiens pour la prise de décisions lors d'analyses intermédiaires semblent être des designs suffisamment flexibles et adapter à ces problématiques. La méthodologie vue précédemment pour la sélection de dose, avec des tests fréquentistes à chaque étape de l'essai et le calcul de la puissance prédictive pour la prise de décision aux analyses intermédiaires peut être utilisée [3]. Ici la décision est la sélection de la population la plus sensible à un traitement.

Pour illustrer l'utilisation des outils bayésiens pour la sélection de sous-population, considérons une population entière $\{F\}$ et une sous-population $\{S\}$ de $\{F\}$. La sous-population a été identifiée au préalable lors d'une étude exploratoire de phase II séparée. Dans chacune des populations, nous testons un traitement afin de sélectionner la population la plus sensible au traitement.

Lors de l'analyse intermédiaire, les données accumulées vont permettre de calculer la puissance prédictive qui est la probabilité de rejeter les différentes hypothèses nulles à la fin de l'étude, mais aussi de calculer les probabilités a posteriori qui peuvent être utilisés pour estimer l'effet traitement dans les différentes populations.

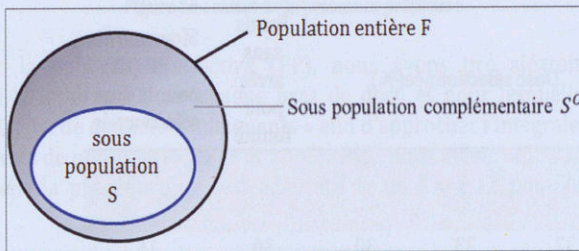


Figure17 : schéma population

La puissance prédictive va permettre, comme dans le cadre de la sélection de dose, de sélectionner la population qui ira à l'étape B ou d'arrêter l'étude précocement pour futilité. Ainsi, la sélection de population grâce à la PP, se fait de la manière suivante : soient les seuils prédéfinis $\pi^{(F)}$ et $\pi^{(S)}$ si :

- $PP^{(F)} \leq \pi^{(F)}$ et $PP^{(S)} \leq \pi^{(S)}$ alors l'étude s'arrête précocement pour futilité. En effet, comme précédemment avec la sélection de dose, si la PP est suffisamment petite, nous pouvons conclure avec une grande confiance que l'essai ne sera pas un succès quelque soit la population sélectionnée.
- $PP^{(F)} \leq \pi^{(F)}$ et $PP^{(S)} > \pi^{(S)}$ alors la sous-population S est sélectionnée. Contrairement à la sélection de dose où la dose avec la PP la plus grande est sélectionnée, ici nous sommes en présence d'une seule sous-population et nous décidons de continuer à l'étape B avec cette sous-population si sa PP dépasse un certain seuil. Ainsi, nous pouvons conclure avec une grande confiance que l'essai sera un succès si des patients de la sous-population sont recrutés pour l'étape B.
- $PP^{(F)} > \pi^{(F)}$ et $PP^{(S)} \leq \pi^{(S)}$ alors la population entière est sélectionnée pour la suite de l'essai. Puisque dans ce cas la PP est suffisamment grande, nous pouvons conclure avec une grande confiance que l'essai sera un succès si des patients de la population entière sont recrutés à l'étape B.

Nous pouvons aussi être dans la situation où les puissances prédictives pour la population entière et la sous-population dépassent les seuils prédéfinis. Dans ce cas, il est légitime de se demander si la PP de la population entière est grande uniquement à cause de l'influence de la sous-population et inversement. Pour répondre à ces interrogations, la probabilité a posteriori, qui estime l'effet traitement, dans la population complémentaire doit être calculée. Dans ce cas, si $PP^{(F)} > \pi^{(F)}$ et $PP^{(S)} > \pi^{(S)}$:

- La sous-population $\{S\}$ est sélectionnée pour la seconde étape de l'essai si $P^{(S^c)} < \pi^{(S^c)}$ où, $P^{(S^c)}$ est la probabilité a posteriori dans la sous-population complémentaire $\{S^c\}$ de $\{S\}$ et $\pi^{(S^c)}$ un seuil prédéfini. Dans ce cas, il n'y a pas d'effet traitement dans la population complémentaire et nous pouvons conclure que $PP^{(F)} > \pi^{(F)}$ uniquement sous l'influence de la sous-population.
- La population entière $\{F\}$ est sélectionnée pour la seconde étape de l'essai si $P^{(S^c)} \geq \pi^{(S^c)}$, car dans ce cas il y a un effet traitement dans la population complémentaire.

Conclusion

Les designs de phases II/III, dans le cadre de développement d'un traitement en oncologie, sont intéressants car contrairement à un design classique de phase II puis de phase III, les patients recrutés à l'étape A et à l'étape B sont utilisés pour l'évaluation finale. Cela peut donc permettre de planifier un essai avec moins de patients et atteindre le même degré de précision qu'un design classique. De plus, ici, le même critère de jugement est utilisé lors des deux étapes du design, contrairement à un design classique où généralement des critères de jugement différents sont utilisés pour la phase II et pour la phase III ce qui peut conduire à avoir un fort taux d'attrition. En effet, garder le même critère de jugement, va permettre d'éviter les mauvais choix de doses ou de population. C'est un aspect important de ce genre de design et un des atouts pour lesquels ils répondent aux problématiques des nouvelles thérapies en oncologie.

L'outil bayésien, quant à lui, permet une modélisation réaliste et de plus en plus précise, grâce à la mise à jour, en fonction des données disponibles au cours de l'essai, de notre paramètre d'intérêt. Le calcul de puissances prédictives, lors d'analyses intermédiaires, n'influe pas sur l'erreur de type I et permet de planifier au mieux la seconde partie de l'essai. La puissance prédictive va permettre d'arrêter l'essai précocement si nous constatons que la dose de traitement est inefficace ou au contraire efficace. De plus, elle permet de choisir la ou les doses les plus prometteuses ou encore la ou les populations les plus réceptive à un traitement. Ainsi, l'analyse bayésienne répond de manière intuitive à des questions cliniques grâce à la puissance prédictive qui est un outil beaucoup plus informatif et transparent qu'une p-value.

Lors de ce stage, un temps conséquent a été consacré à l'utilisation de l'analyse bayésienne dans le cadre des essais de phase II, puis à la recherche et à l'implémentation d'un design adaptatif de phase II/III seamless. Des questions comme le nombre d'analyses intermédiaires, le moment optimal où les effectuer ou encore les seuils critiques nous permettant de choisir les doses ou les populations n'ont pas été étudiées. De plus, il serait intéressant de comparer un design classique de phase II et III avec un design adaptatif de phase II/III seamless.

Annexes

I. Annexe 1 : Calcul de la distribution a posteriori

D'après le théorème de Bayes :

$$\pi(\theta_E | x_M) = \frac{f(x_M | \theta_E) \cdot \pi(\theta_E)}{m(x_M)}$$

Avec :

- Le nombre de réponse X_M suit une loi binomiale avec une probabilité de succès égale à θ_E :

$$X_M \sim \mathcal{B}(M, \theta_E) \Leftrightarrow f(x_M | \theta_E) = \binom{M}{x_M} \theta_E^{x_M} (1 - \theta_E)^{M-x_M}$$

- $\theta_E \sim \beta(a_E, b_E) \Leftrightarrow \pi(\theta_E) = \frac{1}{B(a_E, b_E)} \theta_E^{a_E-1} (1 - \theta_E)^{b_E-1}$ avec $B(a_E, b_E) = \frac{\Gamma(a_E) \Gamma(b_E)}{\Gamma(a_E + b_E)}$

Ainsi la distribution marginale de X_M est calculée de la façon suivante :

$$\begin{aligned} m(x_M) &= \int \binom{M}{x_M} \theta_E^{x_M} (1 - \theta_E)^{M-x_M} \frac{1}{B(a_E, b_E)} \theta_E^{a_E-1} (1 - \theta_E)^{b_E-1} d\theta_E \\ &= \binom{M}{x_M} \frac{B(x_M + a_E - 1, M - x_M + b_E - 1)}{B(a_E, b_E)} \int \frac{1}{B(x_M + a_E - 1, M - x_M + b_E - 1)} \theta_E^{x_M + a_E - 1} \\ &\quad \cdot (1 - \theta_E)^{M - x_M + b_E - 1} d\theta_E \\ &= \binom{M}{x_M} \frac{B(x_M + a_E - 1, M - x_M + b_E - 1)}{B(a_E, b_E)} \end{aligned}$$

Nous en déduisons directement la distribution a posteriori de θ_E conditionnellement aux observations X_M :

$$\begin{aligned} \pi(\theta_E | x_M) &= \binom{M}{x_M} \theta_E^{x_M} (1 - \theta_E)^{M-x_M} \frac{1}{B(a_E, b_E)} \theta_E^{a_E-1} (1 - \theta_E)^{b_E-1} \\ &\quad \frac{B(a_E, b_E)}{\binom{M}{x_M} B(x_M + a_E - 1, M - x_M + b_E - 1)} \\ &= \frac{\theta_E^{x_M + a_E - 1} (1 - \theta_E)^{M - x_M + b_E - 1}}{B(x_M + a_E - 1, M - x_M + b_E - 1)} \end{aligned}$$

Donc :

$$\theta_E | X_M \sim \beta(a_E + x_M, b_E + M - x_M)$$

$$\mathbb{E}(\theta_E | X_M) = \frac{a_E + x_M}{a_E + b_E + M} \text{ et } \mathbb{V}(\theta_E) = \frac{a_E + x_M}{(a_E + b_E + M)^2}$$

II. Annexe 2 : Algorithme pour la table 3

M est déterminé de la façon suivante :

Pour $M \in [M_{\min} = 10, M_{\max} = 100]$, μ_E, p_l et Δ_l fixés

- Simulation du recrutement de un patient pour incrémenter la taille de l'échantillon : $M = M + 1$
- Calcul de $X_M = \mu_E(a_E + b_E + M) - a_E = \mu_E(2 + M) - 1$
- Calcul des paramètres de la distribution a posteriori de $\theta_E | X_M$:

$$\begin{cases} a_{\theta_E | X_M} = a_E + X_M = 1 + X_M \\ b_{\theta_E | X_M} = a_E + b_E + M = 2 + M \end{cases}$$

- Calcul des quantiles L et U qui dépendent de la précision p_l fixée :

$$\begin{cases} U = q_{0.5+p_l/2}(\beta(a_E + X_M, a_E + b_E + M)) \\ L = q_{0.5-p_l/2}(\beta(a_E + X_M, a_E + b_E + M)) \end{cases}$$

- Si $U-L$ la largeur de l'intervalle est égale à Δ_l la largeur d'intervalle désirée, M est trouvé sinon le recrutement continu.

Code R pour la table 3 :

```
taille <- function(a,b,nmin,nmax,m,w,p){
  Found=0
  n=0
  while(Found==0 & n <= nmax){
    n=n+1
    Xn=round((a+b+n)*(m)-a)
    U <- qbeta(0.5+p/2,a+Xn,b+n-Xn)
    L <- qbeta(0.5-p/2,a+Xn,b+n-Xn)
    delta <- U-L
    Fonc <- function(x){dbeta(x,a+Xn,b+n-Xn)}
    int=integrate(Fonc,L,U,rel.tol=0.0001,abs.tol=0)
    if (delta <= w){Found=1}
  }
  Mat <- matrix(0,1,5)
  colnames(Mat) <- c("taille","ic.lower","ic.upper","width","prob.couv")
  Mat[1,]<-c(n,L,U,delta,int$value)
  return(Mat)
}
```


Pour déterminer la taille de l'échantillon, une autre approche peut être vue. En effet, pour l'approche précédente, on fixe p_l et on en déduit les quantiles U et L permettant de trouver M. On peut aussi fixer U et L de sorte d'avoir largeur d'intervalle Δ_l centrée autour de la moyenne et trouver M telle que $\mathbb{P}(L < \theta_E < U | X_M = x_M) \geq p_l$:

```
nb.patient <- function(a,b,nmin,nmax,m,w,p){
  Found=0
  n=0
  while( Found==0 & n < nmax){
    n=n+1
    Xn=round((a+b+n)*(m)-a)
    U=m+(w/2)
    L=m-(w/2)

    Fonc <- function(x){dbeta(x,a+Xn,b+n-Xn)}
    int=integrate(Fonc,L,U,rel.tol=0.0001,abs.tol=0)

    if(int$value > p){Found=1}
  }
  Mat <- matrix(0,1,5)
  colnames(Mat) <- c("taille","ic.lower","ic.upper","width","prob.couv")
  Mat[1,]<-c(n,L,U,w,int$value)
  return(Mat)
}
```

On retrouve les mêmes résultats avec les deux codes.

III. Annexe 3 : Détail de la chaîne de Markov

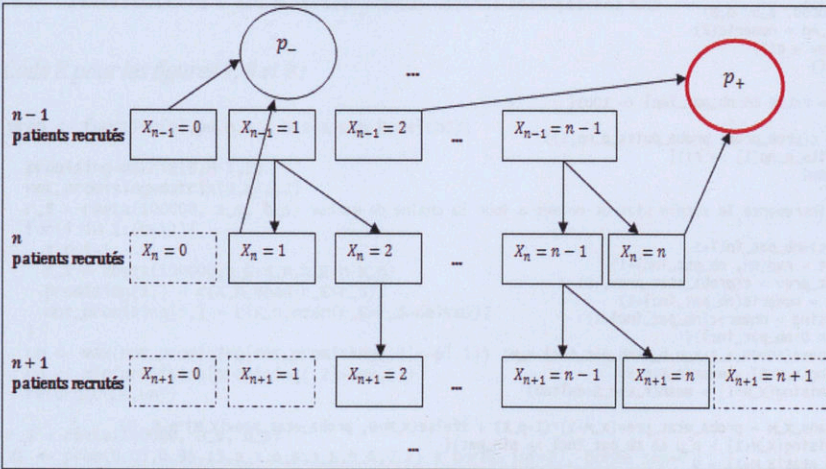


Figure 18 : Détails de la chaîne de Markov permettant de calculer le nombre de sujet nécessaire

Pour n patients recrutés, la probabilité d'être dans les états réponses $X_n = i$, $i \in \{1, \dots, n-1\}$ est calculée de la façon suivante :

$$\begin{aligned} \mathbb{P}(X_n = 0) &= (1 - p_e) \mathbb{P}(X_{n-1} = 0), \text{ où } p_e = \mu_s + \delta_0 \\ \mathbb{P}(X_n = i) &= (1 - p_e) \mathbb{P}(X_{n-1} = i) + p_e \mathbb{P}(X_{n-1} = i - 1) \\ \mathbb{P}(X_n = n) &= p_e \mathbb{P}(X_{n-1} = n - 1) \end{aligned}$$

Si pour une réponse $X_n = i$, le traitement expérimental est prometteur ou non prometteur, la probabilité d'aller dans un des deux états absorbant est mise à jour de la façon suivante :

$$\begin{cases} \mathbb{P}_n(p_{+/-}) = \mathbb{P}_{n-1}(p_{+/-}) + \mathbb{P}(X_n = i) \\ \mathbb{P}(X_{n+1} = i | X_n = i) = \mathbb{P}(X_{n+1} = i + 1 | X_n = i) = 0 \end{cases}$$

Code R pour trouver les paramètres de la distribution du traitement standard :

```
beta_para <- fonction(m,w,p1){
  f <- function(b)((qbeta(0.5+p1/2,b*(m/(1-m)),b)-qbeta(0.5-p1/2,b*(m/(1-m)),b))- w)^2
  b <- optimize(f,c(0,500),tol = 0.001)$minimum
  a <- b*(m/(1-m))
  s <- ((m*(1-m))/(a+b+1))^(1/2)
  return(params = list(alpha = a, beta = b,ecart.type=s))
}
```

Code R pour les tables 4, 5 et la figure 6 :

```
nb_pat = fonction(mu_S, delta0, w_S, pct, Pi, p_U, p_L, min_pat, p_E){
  a_E=1
  b_E=1
  nb_pat_incl = 0
  found = FALSE
  param_S = beta_para(mu_S, w_S,pct)
  a_S = param_S[1]
  b_S = param_S[2]
  r_S = rbeta(50000, a_S, b_S)
  proba_puits_p_np = numeric(2)
  proba_etat_prev = c(1)
  proba_prom = c()

  while(found == FALSE && nb_pat_incl <= 100){
    # p+ >= pi
    proba_prom = c(proba_prom, proba_puits_p_np[1])
    if(proba_puits_p_np[1] >= Pi){
      found = TRUE
    }
    # Sinon on incremente le sample size et on met a jour la chaine de markov
    else{
      nb_pat_incl=nb_pat_incl+1
      proba_etat = rep(NA, nb_pat_incl+1)
      proba_etat_prev = c(proba_etat_prev, 0)
      promising = numeric(nb_pat_incl+1)
      not_promising = numeric(nb_pat_incl+1)
      for(X_M in 0:nb_pat_incl){
        r_E = rbeta(50000,a_E+X_M,b_E+nb_pat_incl-X_M)
        promising[X_M+1] = mean(r_E>r_S)
        not_promising[X_M+1] = mean(r_E>r_S+delta0)

        proba_trans_X_M = proba_etat_prev[X_M+1]*(1-p_E) + ifelse(X_M>0, proba_etat_prev[X_M]*p_E, 0)
        if(promising[X_M+1] > p_U && nb_pat_incl >= min_pat){
          proba_etat[X_M+1] = 0
        }
      }
    }
  }
}
```



```

        proba_puits_p_np[1] = proba_puits_p_np[1] + prob_trans_X_M
    } else if(not_promising[X_M+1] < p_L && nb_pat_incl >= min_pat){
        proba_etat[X_M+1] = 0
        proba_puits_p_np[2] = proba_puits_p_np[2] + prob_trans_X_M
    } else{
        proba_etat[X_M+1] = prob_trans_X_M
    }
    proba_etat_prev = proba_etat
}
}
return(list(found, nb_pat_incl, prob_prom))
}

g1 = nb_pat(0.50, 0.20, 0.30, 0.90, 0.90, 0.95, 0.05, 10, 0.5)
prom1 = g1[[3]]
g2 = nb_pat(0.50, 0.20, 0.30, 0.90, 0.90, 0.95, 0.05, 10, 0.55)
prom2 = g2[[3]]
g3 = nb_pat(0.50, 0.20, 0.30, 0.90, 0.90, 0.95, 0.05, 10, 0.60)
prom3 = g3[[3]]
g4 = nb_pat(0.50, 0.20, 0.30, 0.90, 0.90, 0.95, 0.05, 10, 0.65)
prom4 = g4[[3]]
g5 = nb_pat(0.50, 0.20, 0.30, 0.90, 0.90, 0.95, 0.05, 10, 0.70)
prom5 = g5[[3]]
x_abs = 10:101

plot(x_abs, prom1[10:length(prom1)], ylim=c(0,1), type='s', ylab="P(E declared promising)",
     xlab="Maximum Sample Size")
lines(prom2[10:length(prom2)], type='s')
lines(prom3[10:length(prom3)], type='s')
lines(prom4[10:length(prom4)], type='s')
lines(prom5[10:length(prom5)], type='s')

```

IV. Annexe 4 : Les analyses intermédiaires

Code R pour les figures 7, 8 et 9 :

```

prom <- function(pl,pu,n,a_s,b_s,a_E,b_E,delta0){

    promising=matrix(0,n+1,2)
    not_promising=matrix(0,n+1,2)
    r_s = rbeta(100000, a_s, b_s)
    for(i in 1:(n+1)){
        X_n=i-1
        r_E = rbeta(100000, a_E+X_n, b_E+n-X_n)
        promising[i,] = c(X_n, mean(r_E>r_s))
        not_promising[i,] = c(X_n, mean(r_E>r_s+delta0))
    }
    Ln <- max(not_promising[not_promising[,2]<=pl,1])
    un <- min(promising[promising[,2]>=pu,1])
    return(c(Ln,un))
}

r_s = rbeta(100000, a_s, b_s)
x1 <- prom(0.05,0.95,15,a_s,b_s,a_E,b_E,0.2) # borne inf=2 , borne sup=7
mean(rbeta(100000,a_E+x1[2],b_E+15-x1[2])>r_s)
mean(rbeta(100000,a_E+x1[1],b_E+15-x1[1])>r_s+delta0)

x2 <- prom(0.05,0.95,30,a_s,b_s,a_E,b_E,0.2) # borne inf=6 , borne sup=11
mean(rbeta(100000,a_E+x2[2],b_E+30-x2[2])>r_s)
mean(rbeta(100000,a_E+x2[1],b_E+30-x2[1])>r_s+delta0)

x3 <- prom(0.05,0.95,46,a_s,b_s,a_E,b_E,0.2) # nmax = 46, # borne inf=11 , borne sup=16
mean(rbeta(100000,a_E+x3[2],b_E+46-x3[2])>r_s)
mean(rbeta(100000,a_E+x3[1],b_E+46-x3[1])>r_s+delta0)

```



```

par(mfrow=c(2,2))

p <- seq(0,1,0.01)

plot(p,dbeta(p,a_s,b_s),type="l",col=4,ylab="densité",xlab="taux de réponse",
      ylim=c(0,7))
points(p,dbeta(p,a_E,b_E),type="l",col=2)
legend(0.3,7, c("traitement standard","traitement expérimental"),lty=1
      ,col = c(4,2),bty = "n")
title("début de l'essai")

# nmax=46, AI 1: au bout de 15 patients , AI 2: 15 patients en plus, AI 2: 16 patients en plus

n=15
r_S = rbeta(100000, a_s, b_s)

X_n=4
r_E = rbeta(100000,a_E+X_n,b_E+n-X_n)
plot(p,dbeta(p,a_s,b_s),type="l",col=4,ylab="densité",xlab="taux de réponse",
      ylim=c(0,7))
points(p,dbeta(p,a_E+X_n,b_E+n-X_n),type="l",col=2)
text(0.8,4,paste("efficacité = ",round(mean(r_E>r_S),3),"< 0.95"))
text(0.8,3,paste("futilité = ",round(mean(r_E>r_S+delta0),3),"> 0.05"))
legend(0.3,7, c("traitement standard","traitement expérimental"),lty=1
      ,col = c(4,2),bty = "n")
title("première AI : 15 patients, 4 réponses")

X_n=1
r_E = rbeta(100000,a_E+X_n,b_E+n-X_n)
plot(p,dbeta(p,a_s,b_s),type="l",col=4,ylab="densité",xlab="taux de réponse",
      ylim=c(0,7))
points(p,dbeta(p,a_E+X_n,b_E+n-X_n),type="l",col=2)
text(0.8,4,paste("efficacité = ",round(mean(r_E>r_S),3),"< 0.95"))
text(0.8,3,paste("futilité = ",round(mean(r_E>r_S+delta0),3),"< 0.05"),col="red")
legend(0.3,7, c("traitement standard","traitement expérimental"),lty=1
      ,col = c(4,2),bty = "n")
title("première AI : 15 patients, 1 réponses")

X_n=7
r_E = rbeta(100000,a_E+X_n,b_E+n-X_n)
plot(p,dbeta(p,a_s,b_s),type="l",col=4,ylab="densité",xlab="taux de réponse",
      ylim=c(0,7))
points(p,dbeta(p,a_E+X_n,b_E+n-X_n),type="l",col=2)
text(0.8,4,paste("efficacité = ",round(mean(r_E>r_S),3),"> 0.95"),col="red")
text(0.8,3,paste("futilité = ",round(mean(r_E>r_S+delta0),3),"> 0.05"))
legend(0.3,7, c("traitement standard","traitement expérimental"),lty=1
      ,col = c(4,2),bty = "n")
title("première AI : 15 patients, 7 réponses")

###
X_n=4
r_E = rbeta(100000,a_E+X_n,b_E+n-X_n)
plot(p,dbeta(p,a_s,b_s),type="l",col=4,ylab="densité",xlab="taux de réponse",
      ylim=c(0,7))
points(p,dbeta(p,a_E+X_n,b_E+n-X_n),type="l",col=2)
text(0.8,4,paste("efficacité = ",round(mean(r_E>r_S),3),"< 0.95"))
text(0.8,3,paste("futilité = ",round(mean(r_E>r_S+delta0),3),"> 0.05"))
legend(0.3,7, c("traitement standard","traitement expérimental"),lty=1
      ,col = c(4,2),bty = "n")
title("première AI : 15 patients, 4 réponses")

n=30
X_n=8
r_E = rbeta(100000,a_E+X_n,b_E+n-X_n)
plot(p,dbeta(p,a_s,b_s),type="l",col=4,ylab="densité",xlab="taux de réponse",
      ylim=c(0,7))
points(p,dbeta(p,a_E+X_n,b_E+n-X_n),type="l",col=2)
text(0.8,4,paste("efficacité = ",round(mean(r_E>r_S),3),"< 0.95"))
text(0.8,3,paste("futilité = ",round(mean(r_E>r_S+delta0),3),"> 0.05"))
legend(0.3,7, c("traitement standard","traitement expérimental"),lty=1
      ,col = c(4,2),bty = "n")
title("deuxième AI : 30 patients, 8 réponses")

```



```

X_n=5
r_E = rbeta(100000,a_E+X_n,b_E+n-X_n)
plot(p,dbeta(p,a_s,b_s),type="l",col=4,ylab="Densité",xlab="taux de réponse",
      ylim=c(0,7))
points(p,dbeta(p,a_E+X_n,b_E+n-X_n),type="l",col=2)
text(0.8,4,paste("efficacité = ",round(mean(r_E>r_S),3),"< 0.95"))
text(0.8,3,paste("futilité = ",round(mean(r_E>r_S+delta0),3),"< 0.05"),col="red")
legend(0.3,7, c("traitement standard","traitement expérimental"),lty =1
       ,col = c(4,2),bty = "n")
title("deuxième AI : 30 patients, 5 réponses")

X_n=12
r_E = rbeta(100000,a_E+X_n,b_E+n-X_n)
plot(p,dbeta(p,a_s,b_s),type="l",col=4,ylab="Densité",xlab="taux de réponse",
      ylim=c(0,7))
points(p,dbeta(p,a_E+X_n,b_E+n-X_n),type="l",col=2)
text(0.8,4,paste("efficacité = ",round(mean(r_E>r_S),3),"> 0.95"),col="red")
text(0.8,3,paste("futilité = ",round(mean(r_E>r_S+delta0),3),"> 0.05"))
legend(0.3,7, c("traitement standard","traitement expérimental"),lty =1
       ,col = c(4,2),bty = "n")
title("deuxième AI : 30 patients, 12 réponses")

###
X_n=8
r_E = rbeta(100000,a_E+X_n,b_E+n-X_n)
plot(p,dbeta(p,a_s,b_s),type="l",col=4,ylab="Densité",xlab="taux de réponse",
      ylim=c(0,7))
points(p,dbeta(p,a_E+X_n,b_E+n-X_n),type="l",col=2)
text(0.8,4,paste("efficacité = ",round(mean(r_E>r_S),3),"< 0.95"))
text(0.8,3,paste("futilité = ",round(mean(r_E>r_S+delta0),3),"> 0.05"))
legend(0.3,7, c("traitement standard","traitement expérimental"),lty =1
       ,col = c(4,2),bty = "n")
title("deuxième AI : 30 patients, 8 réponses")

n=46
X_n=13
r_E = rbeta(100000,a_E+X_n,b_E+n-X_n)
plot(p,dbeta(p,a_s,b_s),type="l",col=4,ylab="Densité",xlab="taux de réponse",
      ylim=c(0,7))
points(p,dbeta(p,a_E+X_n,b_E+n-X_n),type="l",col=2)
text(0.8,4,paste("efficacité = ",round(mean(r_E>r_S),3),"< 0.95"))
text(0.8,3,paste("futilité = ",round(mean(r_E>r_S+delta0),3),"> 0.05"))
legend(0.3,7, c("traitement standard","traitement expérimental"),lty =1
       ,col = c(4,2),bty = "n")
title("analyse finale : 46 patients, 13 réponses")

X_n=10
r_E = rbeta(100000,a_E+X_n,b_E+n-X_n)
plot(p,dbeta(p,a_s,b_s),type="l",col=4,ylab="Densité",xlab="taux de réponse",
      ylim=c(0,7))
points(p,dbeta(p,a_E+X_n,b_E+n-X_n),type="l",col=2)
text(0.8,4,paste("efficacité = ",round(mean(r_E>r_S),3),"< 0.95"))
text(0.8,3,paste("futilité = ",round(mean(r_E>r_S+delta0),3),"< 0.05"),col="red")
legend(0.3,7, c("traitement standard","traitement expérimental"),lty =1
       ,col = c(4,2),bty = "n")
title("analyse finale : 46 patients, 10 réponses")

X_n=18
r_E = rbeta(100000,a_E+X_n,b_E+n-X_n)
plot(p,dbeta(p,a_s,b_s),type="l",col=4,ylab="Densité",xlab="taux de réponse",
      ylim=c(0,7))
points(p,dbeta(p,a_E+X_n,b_E+n-X_n),type="l",col=2)
text(0.8,4,paste("efficacité = ",round(mean(r_E>r_S),3),"> 0.95"),col="red")
text(0.8,3,paste("futilité = ",round(mean(r_E>r_S+delta0),3),"> 0.05"))
legend(0.3,8, c("traitement standard","traitement expérimental"),lty =1
       ,col = c(4,2),bty = "n")
title("analyse finale : 46 patients, 18 réponses")

```

Code R de la figure 10 :


```

par(mfrow=c(2,1))
p <- seq(0,1,0.01)
a_E <- 1
b_E <- 1
n <- 46
X_n <- 13

plot(p,dbeta(p,a_s,b_s),type="l",col=4,ylab="Densité",xlab="taux de réponse",
      ylim=c(0,7))
points(p,dbeta(p,a_E,X_n),type="l",col=2)
legend(0.3,7, c("traitement standard","traitement expérimental"),lty=1
      ,col=c(4,2),bty="n")
title("début de l'essai")
dens <- dbeta(p,a_E+X_n,b_E+n-X_n)
plot(p,dbeta(p,a_s,b_s),type="l",col=4,ylab="Densité",xlab="taux de réponse",
      ylim=c(0,7))
points(p,dens,type="l",col=2)
x.shade.up <- seq(0.2,1,0.001)
polygon(c(x.shade.up,rev(x.shade.up)),
        c(dbeta(x.shade.up,a_E+X_n,b_E+n-X_n),0*dbeta(rev(x.shade.up),
        a_E+X_n,b_E+n-X_n)),
        border="red",col="orange")
abline(v=0.28,lty=2)
text(0.8,6,paste("U = ",round(integrate(x,0.4,1)$value,3)))
text(0.8,5,paste("L = ",round(integrate(x,0,0.2)$value,3)))
text(0.8,4,paste("efficacité = ",round(mean(r_E>r_S),2),"< 0.95"))
text(0.8,3,paste("futilité = ",round(mean(r_E>r_S+delta0),2),"> 0.05"))
text(0.8,2,paste("proba(E>0.2) = ",round(mean(r_E>delta0),2),col="red"))
title(main="fin de l'essai",
      sub=paste("nb patients =",n," et nb réponses =",X_n))

```

V. Annexe 5 : Calcul de la probabilité prédictive

Soient y le nombre de réponses pour m futurs patients et x le nombre de réponses pour les n patients observés. Soient $M = n + m$ le nombre total de patients, $z = x + y$ le nombre total de réponses lors de l'analyse finale et $B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$. Pour calculer la probabilité prédictive d'avoir exactement y succès sur m , nous procédons de la manière suivante :

$$\mathbb{P}(y|x) = \int_0^1 \frac{1}{B(a+x, b+n-x)} \theta^{a+x-1} (1-\theta)^{n-x+b-1} \times \binom{m}{y} \theta^y (1-\theta)^{m-y} d\theta$$

$$\mathbb{P}(y|x) = \frac{1}{B(a+x, b+n-x)} \binom{m}{y} \int_0^1 \theta^{a+x+y-1} (1-\theta)^{n+m-(x+y)+b-1} d\theta$$

$$\mathbb{P}(y|x) = \binom{m}{y} \frac{B(a+z, M-z+b)}{B(a+x, b+n-x)} \int_0^1 \frac{1}{B(a+z, M-z+b)} \theta^{a+x+y-1}$$

$$\mathbb{P}(y|x) = \binom{M-n}{y} \frac{B(a+z, M-z+b)}{B(a+x, b+n-x)} \quad (\text{Distribution beta-binomiale})$$

Ainsi la probabilité prédictive d'avoir au moins k réponds sur les m futurs patients à l'analyse finale se calcul comme il suit :

$$\mathbb{P}(y|x) = \sum_{i=k}^m \mathbb{P}(y \geq i|x) = \sum_{i=k}^m \binom{M-n}{i} \frac{B(a+x+i, M-x+i+b)}{B(a+x, b+n-x)}$$

Code R pour le calcul de la probabilité prédictive :


```
# 8 patients supplémentaires pour conclure positivement
library(LearnBayes)
mu_s <- 0.2
a_s <- beta.para(mu_s,0.2,0.9)$alpha
b_s <- beta.para(mu_s,0.2,0.9)$beta
r_s = rbeta(100000, a_s, b_s)
a_E <- 1
b_E <- 1
n <- 46
x1 <- prom(0.05,0.95,n,a_s,b_s,a_E,b_E,0.2) # borne inf=11 , borne sup=16
mean(rbeta(100000,a_E+x1[2],b_E+n-x1[2])>r_s)
mean(rbeta(100000,a_E+x1[1],b_E+n-x1[1])>r_s+delta0)

round(pbetap(c(a_E+8,b_E+30-8),16,8),3)
round(sum(pbetap(c(a_E+8,b_E+30-8),16,8:16)),3)
```

VI. Annexe 6 : Calcul des taux de hasard a posteriori

Soit la fonction de vraisemblance :

$$L(\theta, \text{obs}) = \prod_{i=1}^n f(t_i, \theta)^{c_i} S(t_i, \theta)^{1-c_i}, \text{ où } c_i \text{ est l'indicateur de censure}$$

Soit la distribution a priori du paramètre défini par la densité :

$$\pi(\theta) \propto \theta^{\alpha_0-1} e^{-\beta_0 \theta}$$

Nous en déduisons la distribution a posteriori :

$$\begin{aligned} \pi(\theta|\text{obs}) &\propto L(\theta, \text{obs})\pi(\theta) \\ \pi(\theta|\text{obs}) &\propto \theta^{\sum_{i=0}^n c_i} e^{-\theta \sum_{i=1}^n t_i} \cdot \theta^{\alpha_0-1} e^{-\beta_0 \theta} \\ \pi(\theta|\text{obs}) &\propto \theta^{\sum_{i=0}^n c_i + \alpha_0 - 1} \cdot e^{-\theta(\sum_{i=0}^n t_i + \beta_0)} \end{aligned}$$

On reconnait alors le noyau de la distribution gamma. La distribution a posteriori du paramètre suit alors une loi gamma $\theta|\text{obs} \sim \mathcal{G}(\sum_{i=0}^n c_i + \alpha_0, \sum_{i=0}^n t_i + \beta_0)$.

VII. Annexe 7 : Code R des tables 7,9, 10 et 11

Table 7:


```

GammaPara <- function(a,b,alpha,qinf,f){
  m <- a/b
  a <- c(1:500) # paramètre a (toutes les valeurs possibles de 1 à 500)
  b <- c(round(a/m)) # b=a/m

  i=1
  q=0
  res <- NULL

  tps <- c(0,0.5,1,3,6)
  IC <- tps[2:5] - tps[1:4]

  while(q<=qinf){
    #print(i)

    if(f==1){
      fdr <- 1- exp(-IC[1]*rgamma(50000,a[i],b[i]))
    }

    if(f==2){
      fdr <- 1- exp(-IC[1]*rgamma(50000,2,6))*exp(-IC[2]*rgamma(50000,a[i],b[i]))
    }

    if(f==3){
      fdr <- 1- exp(-IC[1]*rgamma(50000,2,6))*exp(-IC[2]*rgamma(50000,2,16))*
        exp(-IC[3]*rgamma(100000,a[i],b[i]))
    }

    if(f==4){
      fdr <- 1- exp(-IC[1]*rgamma(50000,2,6))*exp(-IC[2]*rgamma(50000,2,16))*
        exp(-IC[3]*rgamma(100000,2,62))*exp(-IC[4]*rgamma(100000,a[i],b[i]))
    }

    q <- round(quantile(fdr,alpha),3)
    res <- rbind(res,c(a=a[i],b=b[i],quantile=q))

    i <- i+1
  }

  res <- res[order(res[,3]),]
  return(res)
}

GammaPara(9,28,0.05,0.03,1)
GammaPara(9,74,0.05,0.07,2)
GammaPara(9,279,0.05,0.11,3)
GammaPara(9,391,0.05,0.15,4)

```

Table 9 :

```

#####

# simulation des données

# a priori on a pour toutes les doses un taux d'événement cumulé de 30% à 6 mois
CDF <- function(a_1,a_2,a_3,a_4,b_1,b_2,b_3,b_4,alpha){
  tps <- c(0,0.5,1,3,6)
  IC <- tps[2:5] - tps[1:4]

  fdr <- 1- exp(-IC[1]*rgamma(50000,a_1,b_1))*exp(-IC[2]*rgamma(50000,a_2,b_2))*
    exp(-IC[3]*rgamma(50000,a_3,b_3))*exp(-IC[4]*rgamma(50000,a_4,b_4))
  return(list(moyenne=round(mean(fdr)*100),q_inf=round(quantile(fdr,0.05)*100),
    q_sup=round(quantile(fdr,0.95)*100)))
}

CDF(2,2,2,2,6,16,62,87,0.05)
CDF(9,9,9,9,28,74,279,391,0.05)

#####

```



```

# A l'AI : tous les patients sont recrutés et la moitié d'entre eux ont été suivis pendant
# 6 mois (ie censurés à 6 mois)
# L'autre moitié est censurée à moins de 6 mois ou ont subi l'événement
# On suppose que à l'AI on a 36 événements pour le bras contrôle (30%*120)
# pour la dose 1 : 30 événements (25%*120)
# pour la dose 2 : 42 événements (35%*120)
# pour la dose 3 : 48 événements (40%*120)
# on suppose que pour toutes les doses, la moitié des événements ont lieu dans la première
# intervalle de temps et que les autres événements sont équitablement répartis dans les
# 3 autres intervalles de temps

AI <- function(nb_event,nb_event_t1,nb_event_t2,nb_event_t3,nb_event_t4){
  data <- rbind(cbind(rep(0,60),rep(6,60)),
               cbind(rep(1,nb_event_t1),round(runif(nb_event_t1,0.1,0.5),digits=2)),
               cbind(rep(1,nb_event_t2),round(runif(nb_event_t2,0.51,1),digits=2)),
               cbind(rep(1,nb_event_t3),round(runif(nb_event_t3,1.1,3),digits=2)),
               cbind(rep(1,nb_event_t4),round(runif(nb_event_t4,3.1,5.9),digits=2)),
               cbind(rep(0,60-nb_event),round(runif(60-nb_event,0.1,5.9),digits=2)))

  return(data)
}

# dose contrôle : (18 événements dans l'intervalle de temps ]0,0.5]; 6 événements dans
# chacun des autres intervalles de temps)

AI_control <- AI(36,18,6,6,6)

# dose 1: (15 événements dans l'intervalle de temps ]0,0.5]; 5 événements dans
# chacun des autres intervalles de temps)

AI_D1 <- AI(30,15,5,5,5)

# dose 2: (21 événements dans l'intervalle de temps ]0,0.5]; 7 événements dans
# chacun des autres intervalles de temps)

AI_D2 <- AI(42,21,7,7,7)

# dose 3: (24 événements dans l'intervalle de temps ]0,0.5]; 8 événements dans
# chacun des autres intervalles de temps)

AI_D3 <- AI(48,24,8,8,8)

colnames(AI_control) <- colnames(AI_D1) <- colnames(AI_D2) <- colnames(AI_D3) <- c('event',
'tps')

# vérification :
trt <-
sum(trt[,1]==1)
sum(trt[,1]==0 & trt[,2]==6)
mean(trt[,1]==1 & trt[,2]<=0.5)
mean(trt[,1]==1 & trt[,2]>0.5 & trt[,2]<=1)
mean(trt[,1]==1 & trt[,2]<=1)
mean(trt[,1]==1 & trt[,2]>1 & trt[,2]<=3)
mean(trt[,1]==1 & trt[,2]<=3)
mean(trt[,1]==1 & trt[,2]>3 & trt[,2]<=6)
mean(trt[,1]==1 & trt[,2]<=6)

# save(AI_control,file="AI_control.Rdata")
# save(AI_D1,file="AI_D1.Rdata")
# save(AI_D2,file="AI_D2.Rdata")
# save(AI_D3,file="AI_D3.Rdata")

#####

# Une fois qu'on a simulé des jeux de données pour les différentes doses, on peut mettre à
# jour la distribution
# Loi à posteriori :

```



```

loi_post <- function(trt,a,b1,b2,b3,b4){
  n <- nrow(trt)
  theta_1 <- theta_2 <- theta_3 <- theta_4 <- matrix(0,n,2,dimnames=list(c(1:n),
                                                                    c('y','tps'))))

  for(i in 1:n){
    if(trt[i,2]<=0.5){theta_1[i,1] <- trt[i,1]
                     theta_1[i,2] <- trt[i,2]
                     theta_2[i,2] <- 0
                     theta_3[i,2] <- 0
                     theta_4[i,2] <- 0}

    if(trt[i,2]>0.5 && trt[i,2]<=1){theta_1[i,2] <- 0.5
                                   theta_2[i,2] <- trt[i,2]-0.5
                                   theta_2[i,1] <- trt[i,1]
                                   theta_3[i,2] <- 0
                                   theta_4[i,2] <- 0}

    if(trt[i,2]>1 && trt[i,2]<=3){theta_1[i,2] <- 0.5
                                theta_2[i,2] <- 0.5
                                theta_3[i,2] <- trt[i,2]-1
                                theta_3[i,1] <- trt[i,1]
                                theta_4[i,2] <- 0}

    if(trt[i,2]>3 && trt[i,2]<=6){theta_1[i,2] <- 0.5
                                theta_2[i,2] <- 0.5
                                theta_3[i,2] <- 2
                                theta_4[i,2] <- trt[i,2]-3
                                theta_4[i,1] <- trt[i,1]}
  }

  theta <- cbind(theta_1,theta_2,theta_3,theta_4)
  theta_1_post <- c(round(a+sum(theta_1[,1])),round(b1+sum(theta_1[,2])))
  theta_2_post <- c(round(a+sum(theta_2[,1])),round(b2+sum(theta_2[,2])))
  theta_3_post <- c(round(a+sum(theta_3[,1])),round(b3+sum(theta_3[,2])))
  theta_4_post <- c(round(a+sum(theta_4[,1])),round(b4+sum(theta_4[,2])))

  return(list(theta_1_post=theta_1_post,theta_2_post=theta_2_post,
             theta_3_post=theta_3_post,theta_4_post=theta_4_post))
}

theta_D1 <- loi_post(AI_D1,2,6,16,62,87)
CDF(theta_D1$theta_1_post[1],theta_D1$theta_2_post[1],theta_D1$theta_3_post[1],
    theta_D1$theta_4_post[1],theta_D1$theta_1_post[2],theta_D1$theta_2_post[2],
    theta_D1$theta_3_post[2],theta_D1$theta_4_post[2],0.05)
theta_D2 <- loi_post(AI_D2,2,6,16,62,87)
CDF(theta_D2$theta_1_post[1],theta_D2$theta_2_post[1],theta_D2$theta_3_post[1],
    theta_D2$theta_4_post[1],theta_D2$theta_1_post[2],theta_D2$theta_2_post[2],
    theta_D2$theta_3_post[2],theta_D2$theta_4_post[2],0.05)
theta_D3 <- loi_post(AI_D3,2,6,16,62,87)
CDF(theta_D3$theta_1_post[1],theta_D3$theta_2_post[1],theta_D3$theta_3_post[1],
    theta_D3$theta_4_post[1],theta_D3$theta_1_post[2],theta_D3$theta_2_post[2],
    theta_D3$theta_3_post[2],theta_D3$theta_4_post[2],0.05)
theta_control <- loi_post(AI_control,9,28,74,279,391)
CDF(theta_control$theta_1_post[1],theta_control$theta_2_post[1],theta_control$theta_3_post[1],
    theta_control$theta_4_post[1],theta_control$theta_1_post[2],theta_control$theta_2_post[2],
    theta_control$theta_3_post[2],theta_control$theta_4_post[2],0.05)

```

Table 10 :


```

# calcul de la puissance prédictive :
# première étape : simuler les événements (ou censures) des patients censurés à un
# temps t < 6 mois ie que serait t-il arrivé si le patient avait été suivi pendant 6 mois?
# deuxième étape : simuler les événements et censures des patients de l'étape B non
# encore recrutés

# on distingue les patients censurés à 6 mois et qui ont subi l'événement de ceux
# qui sont censurés à moins de 6 mois
obs_mis <- function(trt){
  trt_mis <- trt[trt[,1]==0,]
  trt_mis <- trt_mis[trt_mis[,2]<6,]

  trt_obs_1 <- trt[trt[,1]==1,]
  trt_obs_2 <- trt[trt[,2]==6,]
  trt_obs <- rbind(trt_obs_1,trt_obs_2)

  return(list(trt_obs=trt_obs,trt_mis=trt_mis))
}

D1_obs <- obs_mis(AI_D1)$trt_obs
D1_mis <- obs_mis(AI_D1)$trt_mis

D2_obs <- obs_mis(AI_D2)$trt_obs
D2_mis <- obs_mis(AI_D2)$trt_mis

D3_obs <- obs_mis(AI_D3)$trt_obs
D3_mis <- obs_mis(AI_D3)$trt_mis

control_obs <- obs_mis(AI_control)$trt_obs
control_mis <- obs_mis(AI_control)$trt_mis

# simulations des jeux de données manquants
missing_AB <- function(trt_obs,trt_mis,a_1,a_2,a_3,a_4,b_1,b_2,b_3,b_4,M,N){
  tps <- trt_mis[,2]

  y_obs <- matrix(rep(trt_obs[,1],N),nrow=length(trt_obs[,1]),ncol=N)
  y <- NULL

  # Intervalles de temps :
  time <- c(0,0.5,1,3,6)
  DT <- time[2:5]-time[1:4]

  for( k in 1 : M){
    print(k)

    y_mis_A <- matrix(0,length(tps),N)
    y_mis_B <- matrix(0,120,N)

    # la fonction de survie au temps max a posteriori :
    theta_1 <- rgamma(1,a_1,b_1)
    theta_2 <- rgamma(1,a_2,b_2)
    theta_3 <- rgamma(1,a_3,b_3)
    theta_4 <- rgamma(1,a_4,b_4)
    Surv <- exp(-DT[1]*theta_1)*exp(-DT[2]*theta_2)*exp(-DT[3]*theta_3)*exp(-DT[4]*theta_4)

    for(j in 1 : N){
      # Simulation des observations manquantes de l'étape A:

```



```

for(i in 1 : length(tps)){
  if(tps[i]>0 & tps[i]<=0.5){y_mis_A[i,j]=rbinom(1,1,exp(-tps[i]*theta_1)-Surv)}

  if(tps[i]>0.5 & tps[i]<=1){y_mis_A[i,j]=rbinom(1,1,(exp(-0.5*theta_1))*
    (exp(-(tps[i]-0.5)*theta_2))-Surv)}

  if(tps[i]>1 & tps[i]<=3){y_mis_A[i,j]=rbinom(1,1,(exp(-0.5*theta_1))*
    (exp(-0.5*theta_2))*
    (exp(-(tps[i]-1)*theta_3))-Surv)}

  if(tps[i]>3 & tps[i]<=6){y_mis_A[i,j]=rbinom(1,1,(exp(-0.5*theta_1))*
    (exp(-0.5*theta_2))*(exp(-2*theta_3))*
    (exp(-(tps[i]-3)*theta_4))-Surv)}
}

# simulation des observations manquantes de l'étape B:
y_mis_B[,j]=rbinom(120,1,1-Surv)
}

y[[k]] <- rbind(y_obs,y_mis_A,y_mis_B)
}

return(y)
}

M <- 50# nombre de theta tiré
N <- 500# nombre de jeux de données simulé

y_D1 <- missing_AB(D1_obs,D1_mis,theta_D1$theta_1_post[1],theta_D1$theta_2_post[1],
  theta_D1$theta_3_post[1],theta_D1$theta_4_post[1],theta_D1$theta_1_post[2],
  theta_D1$theta_2_post[2],theta_D1$theta_3_post[2],theta_D1$theta_4_post[2],
  M,N)

y_D2 <- missing_AB(D2_obs,D2_mis,theta_D2$theta_1_post[1],theta_D2$theta_2_post[1],
  theta_D2$theta_3_post[1],theta_D2$theta_4_post[1],theta_D2$theta_1_post[2],
  theta_D2$theta_2_post[2],theta_D2$theta_3_post[2],theta_D2$theta_4_post[2],
  M,N)

y_D3 <- missing_AB(D3_obs,D3_mis,theta_D3$theta_1_post[1],theta_D3$theta_2_post[1],
  theta_D3$theta_3_post[1],theta_D3$theta_4_post[1],theta_D3$theta_1_post[2],
  theta_D3$theta_2_post[2],theta_D3$theta_3_post[2],theta_D3$theta_4_post[2],
  M,N)

y_control <- missing_AB(control_obs,control_mis,theta_control$theta_1_post[1],
  theta_control$theta_2_post[1],theta_control$theta_3_post[1],
  theta_control$theta_4_post[1],theta_control$theta_1_post[2],
  theta_control$theta_2_post[2],theta_control$theta_3_post[2],
  theta_control$theta_4_post[2],M,N)

# save(y_control,file="full_control.Rdata")
# save(y_D1,file="full_D1.Rdata")
# save(y_D2,file="full_D2.Rdata")
# save(y_D3,file="full_D3.Rdata")

```



```

#####
##### PROCEDURE DE TEST FERME #####
#####

# armsel=1 -> bras contrôle
# armsel=2 -> traitement 1
# armsel=3 -> traitement 2
# armsel=4 -> traitement 3

test <- function(rep_1_A,rep_2_A,rep_3_A,rep_cont_A,rep_1_B,rep_2_B,rep_3_B,rep_cont_B,armssel,
                 nidelta,alpha){

  if(armssel==2){bras_sel=rep_1_B}
  if(armssel==3){bras_sel=rep_2_B}
  if(armssel==4){bras_sel=rep_3_B}

  # Nb de patients dans chaque bras :
  # Etape A:
  nA <- length(rep_cont_A)

  # Etape A:
  nB <- length(rep_cont_B)

  # Nombre d'événements dans chaque bras de traitement :
  # Etape A:
  RA <- c(sum(rep_cont_A),sum(rep_1_A),sum(rep_2_A),sum(rep_3_A))

  # Etape B:
  RB <- rep(0,4)
  RB[1] <- sum(rep_cont_B)
  RB[armssel] <- sum(bras_sel)

  # Proportions d'événements dans chaque groupe de traitement :
  # Etape A:
  PA <- RA/nA

  # Etape B:
  PB <- RB/nB

  # P-valeurs pour chaque étape et hypothèse :
  # variance du traitement contrôle pour l'étape A et B :
  var_cont_A <- (PA[1]*(1-PA[1]))/nA
  var_cont_B <- (PB[1]*(1-PB[1]))/nB

  # variance pour les traitements à l'étape A et B :
  var_A <- PA[2:4]*(1-PA[2:4])/nA + var_cont_A
  PNIA = pnorm((PA[2:4] - PA[1] - nidelta)/(sqrt(var_A)))
  PSUA = pnorm((PA[2:4] - PA[1])/(sqrt(var_A)))
  var_B = (PB[armssel]*(1-PB[armssel]))/nB + var_cont_B
  PNIB <- rep(1,3)
  PNIB[armssel-1] = pnorm((PB[armssel] - PB[1] - nidelta)/(sqrt(var_B)))
  PSUB <- rep(1,3)
  PSUB[armssel-1] = pnorm((PB[armssel] - PB[1])/(sqrt(var_B)))
  # test de combinaison

  # Etape A
  p1 = PNIA[1] +1e-16
  p2 = PSUA[1] +1e-16
  p3 = PNIA[2] +1e-16
  p4 = PSUA[2] +1e-16
  p5 = PNIA[3] +1e-16
  p6 = PSUA[3] +1e-16

```



```

p13 = min(1, 2*min(p1, p3))
p14 = min(1, 2*min(p1, p4))
p15 = min(1, 2*min(p1, p5))
p16 = min(1, 2*min(p1, p6))
p23 = min(1, 2*min(p2, p3))
p24 = min(1, 2*min(p2, p4))
p25 = min(1, 2*min(p2, p5))
p26 = min(1, 2*min(p2, p6))
p35 = min(1, 2*min(p3, p5))
p36 = min(1, 2*min(p3, p6))
p45 = min(1, 2*min(p4, p5))
p46 = min(1, 2*min(p4, p6))
p135 = min(1, 3*min(p1, p3, p5))
p136 = min(1, 3*min(p1, p3, p6))
p235 = min(1, 3*min(p2, p3, p5))
p236 = min(1, 3*min(p2, p3, p6))
p145 = min(1, 3*min(p1, p4, p5))
p146 = min(1, 3*min(p1, p4, p6))
p245 = min(1, 3*min(p2, p4, p5))
p246 = min(1, 3*min(p2, p4, p6))

```

```
c = exp(-qchisq((1-alpha),df=4,ncp=0)/2)
```

```

a1 = c/p1
a2 = c/p2
a3 = c/p3
a4 = c/p4
a5 = c/p5
a6 = c/p6
a13 = c/p13
a14 = c/p14
a15 = c/p15
a16 = c/p16
a23 = c/p23
a24 = c/p24
a25 = c/p25
a26 = c/p26
a35 = c/p35
a36 = c/p36
a45 = c/p45
a46 = c/p46
a135 = c/p135
a136 = c/p136
a235 = c/p235
a236 = c/p236
a145 = c/p145
a146 = c/p146
a245 = c/p245
a246 = c/p246

```

```
# Etape B
```

```

q1 = PNIB[1]
q2 = PSUB[1]
q3 = PNIB[2]
q4 = PSUB[2]
q5 = PNIB[3]
q6 = PSUB[3]

```

```
# Test final :
```

```

# REJNI="test de NI du bras sélectionné vs le bras contrôle"
# REJSU="test de SUP du bras sélectionné vs le bras contrôle"

if (armSel==2) {
  REJNI = ( q1 < min( a1, a13, a14, a15, a16, a135, a136, a145, a146 ) ) * 1
  REJSU = ( q2 < min( REJNI, a2, a23, a24, a25, a26, a235, a236, a245, a246 ) ) * 1
}

if (armSel==3){
  REJNI = ( q3 < min( a3, a13, a23, a35, a36, a135, a136, a235, a236 ) ) * 1
  REJSU = ( q4 < min( REJNI, a4, a14, a24, a45, a46, a145, a146, a245, a246 ) ) * 1
}

```



```

if (armse1==4){
  REJNI = ( q5 < min(      a5, a15, a25, a35, a45, a135, a145, a235, a245 ) ) *1
  REJSU = ( q6 < min( REJNI, a6, a16, a26, a36, a46, a136, a146, a236, a246 ) ) *1
}

return(list(REJNI=REJNI,REJSU=REJSU))

}

```

Table 11 :

```

source("AI.R")
source("loi_post.R")
source("CDF.R")

#####

# Scénario 1 :

# A l'AI : tous les patients sont recrutés et la moitié d'entre eux ont été suivis pendant
# 6 mois (ie censurés à 6 mois)
# L'autre moitié est censurée à moins de 6 mois ou ont subi l'événement
# On suppose que à l'AI on a 36 événements pour le bras contrôle (30%*120)
# pour la dose 1, 2 et 3 :50 événements (42%*120)
# on suppose que pour toutes les doses, la moitié des événements ont lieu dans le première
# intervalle de temps et que les autres événements sont équitablement répartis dans les
# 3 autres intervalles de temps

# dose contrôle : (18 événements dans l'intervalle de temps ]0,0.5]; 6 événements dans
# chacun des autres intervalles de temps)
AI_control <- AI(36,18,6,6,6)

# dose 1, 2 et 3: (26 événements dans l'intervalle de temps ]0,0.5]; 8 événements dans
# chacun des autres intervalles de temps)
AI_D1_sc1 <- AI(50,26,8,8,8)
AI_D2_sc1 <- AI(50,26,8,8,8)
AI_D3_sc1 <- AI(50,26,8,8,8)

theta_control <- loi_post(AI_control,9,28,74,279,391)
CDF(theta_control$theta_1_post[1],theta_control$theta_2_post[1],theta_control$theta_3_post[1],
    theta_control$theta_4_post[1],theta_control$theta_1_post[2],theta_control$theta_2_post[2],
    theta_control$theta_3_post[2],theta_control$theta_4_post[2],0.05)

theta_D1_sc1 <- loi_post(AI_D1_sc1,2,6,16,62,87)
CDF(theta_D1_sc1$theta_1_post[1],theta_D1_sc1$theta_2_post[1],theta_D1_sc1$theta_3_post[1],
    theta_D1_sc1$theta_4_post[1],theta_D1_sc1$theta_1_post[2],theta_D1_sc1$theta_2_post[2],
    theta_D1_sc1$theta_3_post[2],theta_D1_sc1$theta_4_post[2],0.05)

theta_D2_sc1 <- loi_post(AI_D2_sc1,2,6,16,62,87)
CDF(theta_D2_sc1$theta_1_post[1],theta_D2_sc1$theta_2_post[1],theta_D2_sc1$theta_3_post[1],
    theta_D2_sc1$theta_4_post[1],theta_D2_sc1$theta_1_post[2],theta_D2_sc1$theta_2_post[2],
    theta_D2_sc1$theta_3_post[2],theta_D2_sc1$theta_4_post[2],0.05)

theta_D3_sc1 <- loi_post(AI_D3_sc1,2,6,16,62,87)
CDF(theta_D3_sc1$theta_1_post[1],theta_D3_sc1$theta_2_post[1],theta_D3_sc1$theta_3_post[1],
    theta_D3_sc1$theta_4_post[1],theta_D3_sc1$theta_1_post[2],theta_D3_sc1$theta_2_post[2],
    theta_D3_sc1$theta_3_post[2],theta_D3_sc1$theta_4_post[2],0.05)

# Scénario 2 :

# dose 1, 2 et 3: (35%*120 ie 42 événements ,21 événements dans l'intervalle de temps
# ]0,0.5]; 7 événements dans chacun des autres intervalles de temps)
AI_D1_sc2 <- AI(42,21,7,7,7)
AI_D2_sc2 <- AI(42,21,7,7,7)
AI_D3_sc2 <- AI(42,21,7,7,7)

theta_D1_sc2 <- loi_post(AI_D1_sc2,2,6,16,62,87)
CDF(theta_D1_sc2$theta_1_post[1],theta_D1_sc2$theta_2_post[1],theta_D1_sc2$theta_3_post[1],
    theta_D1_sc2$theta_4_post[1],theta_D1_sc2$theta_1_post[2],theta_D1_sc2$theta_2_post[2],
    theta_D1_sc2$theta_3_post[2],theta_D1_sc2$theta_4_post[2],0.05)

theta_D2_sc2 <- loi_post(AI_D2_sc2,2,6,16,62,87)
CDF(theta_D2_sc2$theta_1_post[1],theta_D2_sc2$theta_2_post[1],theta_D2_sc2$theta_3_post[1],
    theta_D2_sc2$theta_4_post[1],theta_D2_sc2$theta_1_post[2],theta_D2_sc2$theta_2_post[2],
    theta_D2_sc2$theta_3_post[2],theta_D2_sc2$theta_4_post[2],0.05)

```



```
theta_D3_sc2 <- loi_post(AI_D3_sc2,2,6,16,62,87)
CDF(theta_D3_sc2$theta_1_post[1],theta_D3_sc2$theta_2_post[1],theta_D3_sc2$theta_3_post[1],
  theta_D3_sc2$theta_4_post[1],theta_D3_sc2$theta_1_post[2],theta_D3_sc2$theta_2_post[2],
  theta_D3_sc2$theta_3_post[2],theta_D3_sc2$theta_4_post[2],0.05)
```

```
# Scénario 3 :
```

```
# dose 1, 2 et 3: (30%*120 ie 36 événements ,18 événements dans l'intervalle de temps
# ]0,0.5]; 6 événements dans chacun des autres intervalles de temps)
AI_D1_sc3 <- AI(36,18,6,6,6)
AI_D2_sc3 <- AI(36,18,6,6,6)
AI_D3_sc3 <- AI(36,18,6,6,6)
```

```
theta_D1_sc3 <- loi_post(AI_D1_sc3,2,6,16,62,87)
CDF(theta_D1_sc3$theta_1_post[1],theta_D1_sc3$theta_2_post[1],theta_D1_sc3$theta_3_post[1],
  theta_D1_sc3$theta_4_post[1],theta_D1_sc3$theta_1_post[2],theta_D1_sc3$theta_2_post[2],
  theta_D1_sc3$theta_3_post[2],theta_D1_sc3$theta_4_post[2],0.05)
```

```
theta_D2_sc3 <- loi_post(AI_D2_sc3,2,6,16,62,87)
CDF(theta_D2_sc3$theta_1_post[1],theta_D2_sc3$theta_2_post[1],theta_D2_sc3$theta_3_post[1],
  theta_D2_sc3$theta_4_post[1],theta_D2_sc3$theta_1_post[2],theta_D2_sc3$theta_2_post[2],
  theta_D2_sc3$theta_3_post[2],theta_D2_sc3$theta_4_post[2],0.05)
```

```
theta_D3_sc3 <- loi_post(AI_D3_sc3,2,6,16,62,87)
CDF(theta_D3_sc3$theta_1_post[1],theta_D3_sc3$theta_2_post[1],theta_D3_sc3$theta_3_post[1],
  theta_D3_sc3$theta_4_post[1],theta_D3_sc3$theta_1_post[2],theta_D3_sc3$theta_2_post[2],
  theta_D3_sc3$theta_3_post[2],theta_D3_sc3$theta_4_post[2],0.05)
```

```
# Scénario 4 :
```

```
# dose 1: 20%*120 ie 24 événements (12 événements dans l'intervalle de temps ]0,0.5];
# 4 événements dans chacun des autres intervalles de temps)
# dose 2: 30%*120 ie 36 événements (18 événements dans l'intervalle de temps ]0,0.5];
# 6 événements dans chacun des autres intervalles de temps)
# dose 3: 40%*120 ie 48 événements (24 événements dans l'intervalle de temps ]0,0.5];
# 8 événements dans chacun des autres intervalles de temps)
AI_D1_sc4 <- AI(24,12,4,4,4)
AI_D2_sc4 <- AI(36,18,6,6,6)
AI_D3_sc4 <- AI(48,24,8,8,8)
```

```
theta_D1_sc4 <- loi_post(AI_D1_sc4,2,6,16,62,87)
CDF(theta_D1_sc4$theta_1_post[1],theta_D1_sc4$theta_2_post[1],theta_D1_sc4$theta_3_post[1],
  theta_D1_sc4$theta_4_post[1],theta_D1_sc4$theta_1_post[2],theta_D1_sc4$theta_2_post[2],
  theta_D1_sc4$theta_3_post[2],theta_D1_sc4$theta_4_post[2],0.05)
```

```
theta_D2_sc4 <- loi_post(AI_D2_sc4,2,6,16,62,87)
CDF(theta_D2_sc4$theta_1_post[1],theta_D2_sc4$theta_2_post[1],theta_D2_sc4$theta_3_post[1],
  theta_D2_sc4$theta_4_post[1],theta_D2_sc4$theta_1_post[2],theta_D2_sc4$theta_2_post[2],
  theta_D2_sc4$theta_3_post[2],theta_D2_sc4$theta_4_post[2],0.05)
```

```
theta_D3_sc4 <- loi_post(AI_D3_sc4,2,6,16,62,87)
CDF(theta_D3_sc4$theta_1_post[1],theta_D3_sc4$theta_2_post[1],theta_D3_sc4$theta_3_post[1],
  theta_D3_sc4$theta_4_post[1],theta_D3_sc4$theta_1_post[2],theta_D3_sc4$theta_2_post[2],
  theta_D3_sc4$theta_3_post[2],theta_D3_sc4$theta_4_post[2],0.05)
```


Scénario 5 :

```
# dose 1: 25%*120 ie 30 événements (15 événements dans l'intervalle de temps ]0,0.5];
# 5 événements dans chacun des autres intervalles de temps)
# dose 2: 30%*120 ie 36 événements (18 événements dans l'intervalle de temps ]0,0.5];
# 6 événements dans chacun des autres intervalles de temps)
# dose 1: 35%*120 ie 42 événements (21 événements dans l'intervalle de temps ]0,0.5];
# 7 événements dans chacun des autres intervalles de temps)
AI_D1_sc5 <- AI(30,15,5,5,5)
AI_D2_sc5 <- AI(36,18,6,6,6)
AI_D3_sc5 <- AI(42,21,7,7,7)

theta_D1_sc5 <- loi_post(AI_D1_sc5,2,6,16,62,87)
CDF(theta_D1_sc5$theta_1_post[1],theta_D1_sc5$theta_2_post[1],theta_D1_sc5$theta_3_post[1],
    theta_D1_sc5$theta_4_post[1],theta_D1_sc5$theta_1_post[2],theta_D1_sc5$theta_2_post[2],
    theta_D1_sc5$theta_3_post[2],theta_D1_sc5$theta_4_post[2],0.05)

theta_D2_sc5 <- loi_post(AI_D2_sc5,2,6,16,62,87)
CDF(theta_D2_sc5$theta_1_post[1],theta_D2_sc5$theta_2_post[1],theta_D2_sc5$theta_3_post[1],
    theta_D2_sc5$theta_4_post[1],theta_D2_sc5$theta_1_post[2],theta_D2_sc5$theta_2_post[2],
    theta_D2_sc5$theta_3_post[2],theta_D2_sc5$theta_4_post[2],0.05)

theta_D3_sc5 <- loi_post(AI_D3_sc5,2,6,16,62,87)
CDF(theta_D3_sc5$theta_1_post[1],theta_D3_sc5$theta_2_post[1],theta_D3_sc5$theta_3_post[1],
    theta_D3_sc5$theta_4_post[1],theta_D3_sc5$theta_1_post[2],theta_D3_sc5$theta_2_post[2],
    theta_D3_sc5$theta_3_post[2],theta_D3_sc5$theta_4_post[2],0.05)
```

Scénario 6 :

```
# dose 1: 28%*120 ie 34 événements (17 événements dans l'intervalle de temps ]0,0.5];
# 6 événements dans chacun des autres intervalles de temps)
# dose 2: 32%*120 ie 38 événements (20 événements dans l'intervalle de temps ]0,0.5];
# 6 événements dans chacun des autres intervalles de temps)
# dose 1: 35%*120 ie 42 événements (21 événements dans l'intervalle de temps ]0,0.5];
# 7 événements dans chacun des autres intervalles de temps)
AI_D1_sc6 <- AI(34,16,6,6,6)
AI_D2_sc6 <- AI(38,20,6,6,6)
AI_D3_sc6 <- AI(42,21,7,7,7)

theta_D1_sc6 <- loi_post(AI_D1_sc6,2,6,16,62,87)
CDF(theta_D1_sc6$theta_1_post[1],theta_D1_sc6$theta_2_post[1],theta_D1_sc6$theta_3_post[1],
    theta_D1_sc6$theta_4_post[1],theta_D1_sc6$theta_1_post[2],theta_D1_sc6$theta_2_post[2],
    theta_D1_sc6$theta_3_post[2],theta_D1_sc6$theta_4_post[2],0.05)

theta_D2_sc6 <- loi_post(AI_D2_sc6,2,6,16,62,87)
CDF(theta_D2_sc6$theta_1_post[1],theta_D2_sc6$theta_2_post[1],theta_D2_sc6$theta_3_post[1],
    theta_D2_sc6$theta_4_post[1],theta_D2_sc6$theta_1_post[2],theta_D2_sc6$theta_2_post[2],
    theta_D2_sc6$theta_3_post[2],theta_D2_sc6$theta_4_post[2],0.05)

theta_D3_sc6 <- loi_post(AI_D3_sc6,2,6,16,62,87)
CDF(theta_D3_sc6$theta_1_post[1],theta_D3_sc6$theta_2_post[1],theta_D3_sc6$theta_3_post[1],
    theta_D3_sc6$theta_4_post[1],theta_D3_sc6$theta_1_post[2],theta_D3_sc6$theta_2_post[2],
    theta_D3_sc6$theta_3_post[2],theta_D3_sc6$theta_4_post[2],0.05)
```



```

# M := nombre de theta tiré
# N := nombre de jeux de données simulés pour un theta tiré
# alpha <- 0.025
# nidelta <- 0.12
# armsel <- 2,3 ou 4 ie le bras de dose sélectionné

predictive_power <- function(M,N,alpha,nidelta,armsel){
  Res <- NULL
  NI <- c()
  SUP <- c()
  CP_NI <- c()
  CP_SUP <- c()
  PP <- matrix(0,2,3)
  colnames(PP) <- c("PP", "q_inf", "q_sup")
  rownames(PP) <- c("NI", "SUP")

  for(j in 1 : M){
    # print(j)

    for(i in 1 : N){
      Res[[i]] <- test(y_d1[[j]][1:120,i],y_d2[[j]][1:120,i],y_d3[[j]][1:120,i],
        y_control[[j]][1:120,i],y_d1[[j]][121:240,i],y_d2[[j]][121:240,i],
        y_d3[[j]][121:240,i],y_control[[j]][121:240,i],armsel,nidelta,alpha)

      NI[i] <- Res[[i]]$REJNI
      SUP[i] <- Res[[i]]$REJSU
    }

    CP_NI[j] <- mean(NI)
    CP_SUP[j] <- mean(SUP)
  }

  PP[1,1] <- round(mean(CP_NI),2)
  PP[1,2] <- round(quantile(CP_NI,0.05),2)
  PP[1,3] <- round(quantile(CP_NI,1-0.05),2)

  PP[2,1] <- round(mean(CP_SUP),2)
  PP[2,2] <- round(quantile(CP_SUP,0.05),2)
  PP[2,3] <- round(quantile(CP_SUP,1-0.05),2)

  return(PP)
}

```


Bibliographie

1. Joseph G. Ibrahim, Ming-Hui Chen, Debajyoti Sinha, *Bayesian survival analysis*, Springer series in statistics, 2005
2. Shein-Chung Chow, Mark Chang, Chapman & Hall/CRC Biostatistics Series, *Adaptive design methods in clinical trials*, 2006
3. Werner Brannath and al., *Confirmatory adptive designs with bayesian decision tools for targeted therapy in concology*, Statistics in medecine, vol. 28, 1445-1463, 2009
4. Heinz Schmidli and al., *Bayesian predictive power for interim adaptation in seamless phase II/III trials where the endpoint is survival up to some specified timepoint*, Statistics in Medecine, vol. 26, 4925-4938, 2007
5. Jeff Maca and al., *Adaptive seamless Phase II/III designs: background, operational aspects and examples*, Drug Information Journal, vol. 40, 463-473, 2006
6. E Zuber and al., *Phase II/III seamless adaptive designs with bayesian decision tools for efficient development of targeted therapy in oncology*, université de Vienne, 2006
7. Frank Bretz and al., *Confirmatory seamless phase II/III clinical trials with hypotheses selection at interim : Applications and practical considerations*, Biometrical Journal, vol. 48, 635-643, 2006
8. Frank Bretz and al., *Confirmatory seamless phase II/III clinical trials with hypotheses selection at interim : General concepts*, Biometrical Journal, vol. 48, 623-634, 2006
9. Peter K. Kimani and al., *Practical guidelines for adaptive seamless phase II/III clinical trials that use bayesian methods*, Statistics in Medicine, vol. 28, 917-936, 2012
10. Donald A. Berry , *Adaptive clinical trials in oncology*, Nature Reviews, vol. 9, 199-207, 2012
11. Peter F. Thall and Richard Simon, *A bayesian approach to establishing sample size and monitoring criteria for phase II clinical trials*, Controlled Clinical Trials, Elsevier Science Inc., vol. 15, 463-481, 1994
12. Peter F. Thall et Richard Simon, *Practical Bayesian Guidelines for Phase IIB Clinical Trials*, Biometrics, vol. 50, 337-349, 1994

13. FDA Draft Guidance for Industry: Enrichment Strategies for Clinical Trials, December 2012
14. Peter F. Thall, Donald A. Berry et Lurdes Y. T. Inoue, *Seamlessly Expanding a Randomized Phase II Trial to Phase III*, *Biometrics*, vol. 58, 823-831, 2002
15. Mark Chang et al., *Adaptive design in clinical research - Issues, opportunities, and recommendations*, *Journal of Biopharmaceutical Statistics*, vol. 16, 299-309, 2006
16. Paul Gallo et al., *Adaptive design in clinical drug development - An executive summary of the PhRMA Working Group (with discussions)*, *Journal of Biopharmaceutical Statistics*, vol. 16, 275-283, 2006
17. Raphaël Porcher, cours « essais adaptatifs » et « essais séquentiels », Université Paris Diderot.

INSTRUCTIONS POUR LES AUTEURS

1. Les manuscrits doivent être envoyés par e-mail à la Revue, au format PDF par E-mail : hydra@math.cnrs.fr.

Préparer un fichier PDF :

• 1 page par article

• 10 pages pour le résumé

• 1 page pour la table des matières

• 10 pages pour la table des matières

Un accusé de réception sera adressé aux auteurs.

2. Le style de présentation, titres, références, etc., doit suivre les « instructions aux auteurs » des *Annals of Probability* ou *Annals of Statistics*.

Le groupe des manuscrits sera affecté de préférence par un système de classement de texte ou *latex*.

Le format doit respecter un cadre de 14 cm x 22 cm.

En cas d'acceptation définitive, les auteurs se voient reproduire directement en offset, un fichier au format *pdf* (par exemple *pdf*) sans autre altération que le noir sur un papier blanc.

3. Les manuscrits seront des documents électroniques.
4. Le délai de réponse aux auteurs sera de 10 à 15 jours après soumission.

5. 25 euros à part seront envoyés aux auteurs dont les articles sont publiés.

13. FDA Draft Guidance for Industry: Enrollment Strategies for Clinical Trials, December 2012
14. Peter F. Thall, Donald A. Berry et Lander T. F. Inoue, *Scientifically Expanding a Randomized Phase II Trial to Phase II/III Biomarkers*, vol. 53, 823-831, 2007
15. Mark Gung et al., *Adaptive design in clinical research - Issues, opportunities and recommendations*, Journal of Biopharmaceutical Statistics, vol. 16, 299-309, 2006
16. Paul Gelle et al., *Adaptive design in clinical drug development - An executive summary of the Pfizer Working Group (with discussion)*, Journal of Biopharmaceutical Statistics, vol. 16, 275-283, 2006
17. R  miel Pichier, cours « essais adaptatifs » et « essais s  quentiels », Universit   Paris Diderot