



HAL
open science

Méthodes d'aide à la décision thérapeutique dans les cas des maladies rares : intérêt des méthodes bayésiennes et application à la maladie de Horton

Paméla El Hajj

► To cite this version:

Paméla El Hajj. Méthodes d'aide à la décision thérapeutique dans les cas des maladies rares : intérêt des méthodes bayésiennes et application à la maladie de Horton. *Méthodologie [stat.ME]*. Université Montpellier, 2017. Français. NNT : 2017MONT037 . tel-01921936

HAL Id: tel-01921936

<https://theses.hal.science/tel-01921936>

Submitted on 14 Nov 2018

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.

THÈSE

Pour obtenir le grade de
Docteur

Délivré par l'Université de Montpellier

Préparée au sein de de l'école doctorale n°166 :
Information, Structures, Systèmes (I2S)

Et de l'unité de recherche UPRES EA 2415 : Laboratoire
de Biostatistiques, d'Epidémiologie et de Santé Publique

Spécialité : **Biostatistiques**

Présentée par : **Paméla El Hajj**

**Méthodes d'aide à la décision
thérapeutique dans les cas des
maladies rares : intérêt des
méthodes bayésiennes et application
à la maladie de Horton**

Soutenue le 29 Septembre 2017 devant le jury composé de :

| | | | |
|--------------------|-------|---|--------------|
| Christine LASSET | PU-PH | Université Claude Bernard Lyon 1 | Rapporteur |
| Roland SAMBUC | PU-PH | Hôpital de la Timone Marseille | Rapporteur |
| Serge AMSELEM | PU-PH | Université Pierre et Marie Curie, Paris | Examineur |
| Gérard DURU | PU | Université Claude Bernard Lyon 1 | Invité |
| Jean-Pierre DAURÈS | PU-PH | Université de Montpellier | Directeur |
| Paul LANDAIS | PU-PH | Université de Montpellier | Co-directeur |



Remerciements

A l'issue de la rédaction de cette recherche, je suis convaincue que la thèse est loin d'être un travail solitaire. En effet, je n'aurais jamais pu réaliser ce travail doctoral sans le soutien d'un grand nombre de personnes dont la générosité, la bonne humeur et l'intérêt manifestés à l'égard de ma recherche m'ont permis de progresser dans cette phase délicate de "PhD student".

J'exprime tout d'abord ma profonde reconnaissance au **Pr. Jean-Pierre Daurès**, pour m'avoir supervisé et suivi tout au long de ces années de thèse. Vos compétences et votre expérience ont été indispensables dans l'orientation de ce travail. Vous avez été toujours été disponible, à l'écoute de mes nombreuses questions, et vous vous êtes toujours intéressé à l'avancée de mes travaux. Les nombreuses discussions que nous avons eues ainsi que vos conseils sont pour beaucoup dans le résultat final de ce travail. Enfin, j'ai été extrêmement sensible à ses qualités humaines d'écoute et de compréhension tout au long de ce travail doctoral. Pour tout cela merci.

Je tiens à remercier chaleureusement le **Pr. Paul Landais** d'avoir co-encadré ce travail de thèse. Merci donc pour votre gentillesse, pour vos précieux conseils et pour la motivation et la positivité que vous m'avez fournis surtout les derniers mois qui ont été un peu difficiles. Enfin, vos nombreuses relectures et vos corrections de la thèse et de la présentation ont été très appréciées.

Je remercie le **Pr. Christine Lassetet** et le **Pr. Roland Sambuc** d'avoir accepté de relire cette thèse et d'en être rapporteurs. La version finale de ce mémoire a bénéficié de leur lecture très attentive et de leurs remarques précieuses. Je remercie également le **Pr. Serge Amselem** et le **Pr. Gérard Duru** d'avoir accepté d'assister à la présentation

de ce travail.

Je remercie bien tous mes collègues de l'IURC (Institut Universitaire de Recherche Clinique) pour l'ambiance sympathique. Je ne peux malheureusement pas citer tout le monde alors je me contenterai de faire quelques mentions particulières. Je commence donc par remercier **Sylvie Barthelemy** pour sa gentillesse, son écoute et son efficacité. Je tiens à remercier particulièrement **Amandine Coffy** et **Mohamed Belkacemi** pour leur bonne humeur, leur sympathies et les discussions sérieuses et rigolotes que nous avons eu surtout la dernière année et à **Francoise Perriard** pour son soutien et son encouragement tout au long de ces années de thèse, et enfin je remercie Thierry Pujol pour tous les dépannages informatiques.

Les quatre années au sein de l'IURC sont passées agréablement grâce à trois personnes qui avant ces 4 ans n'étaient pour moi que des étrangers et aujourd'hui ils sont devenus ma famille occidentale : Audrey, Charles, Nathalie (par ordre alphabétique, lol).

Audrey, tes tendres sourires, ta gentillesse et ta sympathie ont ajouté de la gaieté à notre club. Je suis très heureuse d'avoir partagé avec toi le plus beau jour de ta vie, ton mariage surtout avec le "Congratulations and Celebrations". Je suis contente que nos thèses se déroulent dans la même période, car ca nous a permis d'être ensemble moralement, encouragement et organisation. Je te remercie pour tous les beaux moments que nous avons passés ensemble durant ces trois ans au labo et en dehors du labo et surtout pour les merveilleux souvenirs durant nos voyages lors des congrès aux Pays-Bas et en Angleterre. Je suis très contente que la dernière période nous D'autre part, je tiens à te féliciter pour tous les progrès que tu as fait pour prendre des selfies réussis, lol, t'as vu qu'il y a des avantages d'avoir une libanaise à ses côtés. Enfin, je souhaite te dire que tu vas beaucoup me manquer cet année lorsque tu partiras aux Etats-Unis mais j'espère qu'on restera en contact et que l'on conservera notre amitié.

Charles, ta présence ces trois années de thèse était indispensable pour créer cette belle ambiance de joie. Tes petites blagues quotidiennes et tout le vocabulaire "enrichissant" que tu m'as appris m'ont manqué depuis que t'es parti. Je te remercie pour tous les encouragements que l'on s'est donné surtout lors des nombreux week-end de travail au labo. Je ne vais jamais oublier nos petites pauses, nos vendredi midi au resto U, nos petites escapades pour une bière et que tu m'as fait découvrir l'Angleterre en une demi-journée, lol. J'aimerais bien te dire que j'ai apprécié ta présence à mes côtés, dans les bons

comme dans les mauvais moments. Charles tu es un une personne ouverte d'esprit et tu es foncièrement gentil, ne change donc jamais. Je suis ravie que la France m'ait donné un vrai ami. Enfin, je te souhaite un futur brillant comme toi et merci d'avoir fait l'effort de venir depuis l'Arabie Saoudite spécialement pour partager ce moment de réussite avec moi.

Nathalie R., tu es la première personne avec laquelle j'ai sympathisé au labo. Je te remercie pour m'avoir invité à Millau pour passer le week-end et de me faire partager des moments au sein de ta famille. Je te remercie également, pour tous les fous rires que j'ai attrapés à cause de tes blagues rigolotes. Notre club des 4 sans toi aurait été moins drôle. D'autre part, je ne peux pas oublier tous les midis où nous mangions ensemble en ayant plein de discussions aussi rigolotes que sérieuses, tous les moments agréables que nous avons passés ensemble en dehors du labo, merci aussi pour ton enthousiasme et énergie contagieuses surtout au sport. Enfin, je souhaite exprimer ma reconnaissance pour toutes les fois où tu m'as encouragé à ne pas abandonner et tous tes conseils bienveillants. Je te remercie d'avoir accepté d'être ma prof de français tout au long de ces années, d'avoir pris le temps de corriger une grande partie de ma thèse et d'avoir été ma meilleure élève en libanaise. J'espère que notre amitié restera forte en toutes conditions et notamment géographiques.

Cette thèse n'aurait pas aboutie sans l'existence d'amitiés de qualité au cours de ces dernières années. **Awaz, Nathalie M., Anas** j'ai perdu mes mots et je ne sais pas par où commencer. Vous n'êtes pas mes amis, ni mes soeurs et mon frère de sang, mais vous êtes mes soeurs et mon frère de coeur et de choix. **Awaz** tu es mon ange gardien, tu étais toujours à mes côtés pour me féliciter dans les beaux moments et m'encourager et me concilier dans les moments difficiles. Je ne peux pas oublier comment t'as pris soin de moi les trois ans de thèse et surtout les derniers mois de rédaction et je t'en suis reconnaissante. **Nathalie** nous étions déjà amis au Liban, mais en France notre amitié a dépassé tous les stades pour devenir une famille, merci du fond de coeur d'avoir partagé ma vie ces années et d'être toujours à mes côtés dans les moments les plus difficiles en gardant ta bonne humeur et ton sourire. **Anas**, depuis trois ans j'avais un seul frère, Larry, maintenant j'en ai deux, Larry et Anas. Je te remercie car tu étais toujours présent pour moi, chaque fois que j'ai le moindre soucis tu étais la première personne que j'appelais et tu étais toujours prêt à m'aider et à me rendre service. Merci à vous trois et à **Zakarea** pour tous les petits voyages que nous avons faites en France et en Europe et l'ambiance agréable de notre groupe.

Enfin, mes derniers remerciements sont pour mes parents, qui ont cru en moi et en mes capacités et m'ont soutenu sans faille afin de réaliser mon rêve de doctorat, j'ai beaucoup de chance d'avoir des parents comme vous et j'espère que vous êtes fiers de moi, je vous aime. Un grand merci à mon frère **Larry** et mes soeurs **Carla** et **Cynthia** pour leurs accueils chaleureux chaque fois que je reviens au Liban pour les vacances et leurs sentiments sincères d'amour. Maman, Papa, Larry votre présence à mes côtés ce jour important me fait sentir comme la plus heureuse fille au monde, merci et bisous.

Résumé

Les maladies rares touchent moins de 1/2000 personnes. Des questions méthodologiques spécifiques sont dues à cette rareté et en particulier en recherche clinique pour les essais cliniques contrôlés randomisés de supériorité ou de non-infériorité.

Notre objectif était tout d'abord d'identifier des approches novatrices pour la recherche thérapeutique qui ont été, ou peuvent être, utilisées afin de surmonter les difficultés méthodologiques inhérentes à l'étude des maladies rares. Notre recherche systématique de la littérature a ciblé des publications relatives aux caractéristiques des différents designs et méthodes statistiques qui ont été utilisées pour étudier les traitements des maladies rares. Nous nous sommes particulièrement intéressés aux méthodes bayésiennes.

Dans le contexte de ces méthodes il faut introduire une loi informative *a priori* sur l'effet inconnu du traitement. La détermination de la loi *a priori* est délicate. Nous avons recherché les méthodes qui permettent de déterminer la loi *a priori* en considérant des informations provenant d'études historiques et/ou de données provenant d'autres études "voisines". Les informations *a priori* peuvent aussi être résumées par la distribution *a priori* dont les paramètres sont déterminés à partir d'opinions d'experts. Leurs avis sont recueillis sur la base d'un questionnaire permettant d'en déduire les paramètres du modèle bayésien.

Nous avons développé une approche d'élicitation des préférences d'experts pour déterminer une loi *a priori* en prenant pour exemple le traitement d'une maladie rare, la maladie de Horton ou artérite à cellules géantes. Dans ce contexte, nous avons décrit un modèle bayésien qui a pour but de tester l'hypothèse de non-infériorité selon laquelle un traitement associant le méthotrexate aux corticostéroïdes n'est pas moins efficace que le traitement de référence par les corticostéroïdes seuls.

Enfin, nous avons proposé une nouvelle approche dite d'épsilon contamination d'une loi *a priori*. Cette méthode est fondée sur le principe de contaminer une loi *a priori* pas entièrement satisfaisante par une série de lois provenant d'informations issues d'autres études concernant la même maladie, un même traitement ou une même population.

Nous présentons les résultats de ces modélisations dans la dernière partie du travail. Les résultats obtenus pour la maladie de Horton sont favorables à la non-infériorité du traitement associant le méthotrexate aux corticostéroïdes par rapport aux corticostéroïdes seuls. Ces résultats nécessiteront d'être validés par d'autres études.

Mots clés : Maladies rares, Méthodes bayésiennes, "Elicitation" des opinions des experts, Essais de non-infériorité .

Abstract

In Europe, a disease or disorder is defined as rare when it affects less than 1 in 2000 citizens. Specific methodological issues are due to this rarity, particularly when clinical research deals with randomized controlled clinical trials of superiority or non-inferiority. Our objective was first to identify innovative approaches to therapeutic research that have been or may be used to overcome the methodological difficulties inherent to the study of rare diseases. Our systematic research in the literature has targeted publications relative to the characteristics of different designs and statistical methods that have been used to study the treatment of rare diseases. We particularly focused our interest on Bayesian methods.

In the context of these methods, it is necessary to introduce an informative prior distribution for the unknown effect of the treatment. Determination of the prior distribution in a Bayesian model is challenging, we looked for the methods of determining the prior including the possibility of considering information from some historical controlled trials and/or data coming from other studies sufficiently close to the subject of interest. The prior information can be summarized by the prior distribution determined from experts' opinions. A questionnaire is submitted to the experts to elicit their opinions in order to deduce the parameters of the Bayesian model.

We have developed an approach to elicit the experts' opinion to determine a prior distribution using as an example the treatment of Horton's disease a Giant Cell Arteritis. In this context, we described a Bayesian model that aims to test the hypothesis of non-inferiority that a treatment associating methotrexate to the corticosteroids is not less effective than the treatment of reference, the corticosteroids alone.

Finally, we proposed a new approach called the epsilon-contamination method, which is based on contaminating an *a priori* not entirely satisfactory by a series of distributions drawn from information on other studies sharing close conditions, treatments or even populations.

We present the results of these models in the latter part of our work.

The results obtained for Horton's disease are in favor of the non-inferiority of the treatment associating methotrexate to corticosteroids compared to corticosteroids alone. These results will have to be confirmed by other studies.

Keywords : Rare diseases, bayesian methods, elicitation of experts' opinions, non-inferiority trials

Introduction

Les maladies rares sont des maladies qui affectent un petit nombre de personnes par rapport à la population générale. Les autorités médicales dans le monde s'appuient généralement sur la faible prévalence d'une maladie et peu de traitements spécifiques à ces maladies. Une maladie est dite rare en Europe lorsqu'elle affecte 1 cas sur 2000, tandis qu'aux États-Unis elle est définie pour moins de 200000 Américains affectés. Il existe actuellement entre 6000 et 7000 maladies rares, dont 80 % sont d'origine génétique.

Il est parfois difficile de donner une définition commune internationalement de rareté pour une maladie car une maladie peut être rare dans un pays et fréquente dans un autre, ce qui est le cas, par exemple, de la thalassémie qui est une maladie rare en Europe du Nord, mais elle est fréquente dans la région méditerranéenne.

Bien que 6% de la population mondiale ait une maladie rare, la recherche interventionnelle dans ces maladies est limitée et parfois de mauvaise qualité. Le recrutement et le suivi des participants dans des cohortes, en particulier lorsque les maladies sont progressives ou dégénératives, peut être difficile. De ce fait, la recherche clinique dans les maladies rares et les autres populations de petites tailles présente une série de défis statistiques car les grands essais cliniques généralement utilisés pour évaluer l'efficacité de nouveaux médicaments et d'autres interventions de santé sont irréalisables. Le défi le plus évident pour mener des recherches rigoureuses est le problème de puissance statistique réduite dans les études de petite taille car un essai peu puissant n'a généralement pas d'intérêt car il a peu de chance de mettre en évidence l'effet du traitement.

D'un point de vue statistique, certains de ces problèmes peuvent être surmontés en utilisant des modèles d'étude spécifiques et des méthodes statistiques pour l'analyse de données, qui visent généralement à augmenter la puissance de l'étude. Donc l'objectif de ma thèse est de répondre à plusieurs questions scientifiques pour relever des défis statistiques spécifiques aux maladies et aux essais cliniques de petite taille.

Dans ce travail nous avons étudié les différents concepts et méthodes présentés dans la littérature pour répondre à notre problématique, à partir de cette recherche nous sommes basés sur le choix de l'approche bayésienne pour les études des maladies rares. Il s'agit de proposer une terminologie et des définitions claires et cohérentes de l'approche bayésienne et de montrer comment cette approche peut augmenter la robustesse de l'information en utilisant toutes les données disponibles des essais cliniques et d'autres sources "voisines" à fin de fournir une justification pour des tels essais.

Le présent mémoire est organisé en quatre chapitres :

Le premier chapitre est une revue de la littérature afin d'étudier la conception et les méthodes des études récemment publiées dans le cadre des maladies rares, ainsi que la littérature statistique et les lignes directrices sur les études sur les maladies rares ou les études ne disposant que d'un petit nombre d'observations. Dans le cadre d'une première tentative, tous les articles de recherche publiés relatives aux "maladies Rares" ou avec "Petit échantillon" ont été examinés. Beaucoup d'entre eux sont des études de cas ou comprennent uniquement des statistiques descriptives. Dans presque tous les articles trouvés, des méthodes statistiques standards sont utilisées pour l'analyse des données. Dans une deuxième partie, des documents méthodologiques ont été examinés. Souvent, on peut constater qu'il n'y a pas de modèles spécifiquement développés pour les études sur les maladies rares. Toutes les approches qui sont recommandées pour de petits échantillons, dérivent généralement des méthodologies développées pour de grands échantillons, mais elles sont généralement plus complexes, comme par exemple certains modèles adaptatifs ou certains essais croisés. Des modifications de l'application de méthodes standards des échantillons de petite taille sont disponibles dans la littérature mais elles ne sont pas fréquentes.

Le deuxième chapitre illustre les différentes définitions et notations liées à l'approche bayésienne dont l'utilisation est recommandée dans la littérature pour les études des maladies rares, cette approche sera utilisée tout au long du présent mémoire. La suite de ce chapitre est dédiée à un état de l'art sur les méthodes bayésiennes dans le cas des maladies rares. La difficulté que présente l'approche bayésienne concerne surtout la façon d'introduire une loi informative *a priori* sur l'effet inconnu du traitement. Enfin, nous constatons une utilisation croissante des méthodes statistiques bayésiennes en recherche clinique. En effet, l'approche bayésienne est appropriée pour prendre en compte l'information accumulée au cours d'un essai, ce qui permet de rendre informatifs des essais de

taille restreinte et de disposer d'arguments pertinents pour la prise de décision.

Le troisième chapitre présente les méthodes destinées à obtenir l'information *a priori* à partir des connaissances des experts. Dans ce chapitre nous présentons les principales étapes requises par le statisticien pour le design de l'"élicitation" des connaissances des experts, et les méthodes requises pour formaliser les informations de ces experts : l'agrégation mathématique et comportementale qui sont deux approches utilisés pour combiner les différentes informations *a priori* en un seul consensus d'*a priori*.

Le dernier chapitre présente une approche bayésienne pour la conduite d'essais de maladies rares comparant le méthotrexate aux corticostéroïdes pour l'artérite à cellules géantes ou maladie de Horton. Pour étudier l'efficacité du méthotrexate nous présentons trois méthodes différentes nouvelles pour la plupart. La première méthode est un schéma bayésien proposé par [Hampson et al., 2014] dans lequel on calcule les paramètres des lois *a priori* en utilisant les réponses des experts à des questions précises. La deuxième méthode utilise la différence de deux lois bêta, telle que les paramètres de ces lois bêta soient aussi estimés à partir des avis des experts. Finalement, nous utilisons la méthode ϵ -contamination, qui est une approche bayésienne qui jusqu'à présent n'a pas été utilisée dans la littérature pour la conduite des essais de maladies rares comparant un traitement expérimental à un traitement de référence. Cette approche est fondée sur le principe de la contamination d'une loi *a priori* pas entièrement satisfaisante par une série de lois provenant d'informations recueillies à partir d'autres études "proches".

Table des matières

Chapitre 1 – Les méthodes de gestion des maladies rares et des essais cliniques sur petits échantillons..... 19

| | | |
|-------|---|----|
| 1.1 | Introduction | 19 |
| 1.2 | Les méthodes non bayésiennes utilisées dans la littérature | 22 |
| 1.2.1 | Méthodes de recherche | 22 |
| 1.2.2 | Choix des mots clés | 22 |
| 1.2.3 | Recherche dans les bases de données. | 22 |
| 1.2.4 | Revue de la littérature | 22 |
| 1.2.5 | Extraction des articles | 24 |
| 1.3 | Les conceptions d'étude | 24 |
| 1.3.1 | Modèle de démarrage retardé | 24 |
| 1.3.2 | Les essais croisés | 26 |
| 1.3.3 | Les essais de type N-de-1 | 30 |
| 1.3.4 | Plan factoriel | 31 |
| 1.3.5 | Design de retrait aléatoire et modèle d'échappement précoce | 32 |
| 1.3.6 | Conception randomisée de la phase placebo | 33 |
| 1.4 | Modèles adaptatifs | 35 |
| 1.4.1 | Types de modèles adaptatifs | 35 |
| 1.4.2 | Les modèles de randomisation adaptative | 36 |
| 1.4.3 | Formulation mathématique des essais randomisés | 36 |
| 1.5 | Conclusion sur les méthodes adaptatives | 42 |
| 1.6 | Présentation de stratégies statistiques participant à la résolution des difficultés statistiques dans le cas des maladies rares | 43 |
| 1.6.1 | Les difficultés statistiques dans le cas des maladies rares | 43 |
| 1.6.2 | Utilisation des résultats de substitution dans les études des maladies rares | 44 |

| | | |
|--|--|----|
| 1.6.3 | Utilisation des scores de Propension | 44 |
| 1.6.4 | Planifier une méta-analyse prospective | 45 |
| 1.6.5 | Méthodes bayésiennes | 47 |
| 1.7 | Discussion | 49 |
| 1.8 | Avantages et Inconvénients des méthodes | 50 |
| Chapitre 2 – Introduction aux méthodes bayésiennes et leurs applica- tions aux maladies rares | | |
| 2.1 | Introduction | 51 |
| 2.1.1 | Introduction du chapitre | 51 |
| 2.1.2 | Comment se définit la statistique bayésienne | 51 |
| 2.2 | Comment planifier une analyse bayésienne dans les essais cliniques. | 54 |
| 2.3 | L’approche bayésienne | 55 |
| 2.3.1 | L’approche bayésienne | 55 |
| 2.4 | Loi <i>a priori</i> informative et non informative | 62 |
| 2.5 | Modèles hiérarchiques | 64 |
| 2.5.1 | Introduction | 64 |
| 2.5.2 | Théorie | 65 |
| 2.6 | La théorie de décision en analyse bayésienne | 66 |
| 2.7 | Choix du modèle | 70 |
| 2.8 | Critères pour une sélection objective de la loi <i>a priori</i> | 71 |
| 2.9 | Méthodes bayésiennes dans le cas de maladies rares. | 72 |
| 2.10 | Perturbation de la loi <i>a priori</i> par les informations provenant d’autres études notamment chez les adultes. | 72 |
| 2.11 | Méthodes bayésiennes dans le cas des maladies rares | 78 |
| Chapitre 3 – Classement des groupes de traitements et élicitation des avis des experts | | |
| 3.1 | Classement des groupes de traitements | 83 |
| 3.2 | Méthode du tri croisé | 85 |
| 3.2.1 | Introduction aux méthodes de tri croisé | 85 |
| 3.2.2 | Méthode principale aux valeurs propres | 87 |
| 3.2.3 | Méthode basée sur la régression logarithmique selon les moindres carrés ou Méthode des moyennes géométriques | 88 |
| 3.2.4 | Méthode des moindres carrés(M.M.C) | 89 |

| | | |
|--|--|-----|
| 3.2.5 | Cohérence | 91 |
| 3.2.6 | Mesure de la Cohérence | 93 |
| 3.3 | Utilisation des avis d'experts pour trouver l' <i>a priori</i> dans les modèles statistiques bayésiens | 99 |
| 3.3.1 | Questionnaires et élicitation de l'opinion des experts | 99 |
| 3.4 | Prise de décision en présence de plusieurs experts | 102 |
| 3.4.1 | Agrégation mathématique | 102 |
| 3.5 | Agrégation comportementale | 105 |
| 3.6 | Conclusion | 106 |
| Chapitre 4 – Étude de l'efficacité du méthotrexate associé aux corticostéroïdes par rapport à la corticothérapie seule : Résultats des méthodes bayésiennes appliquées | | |
| 4.1 | Introduction | 107 |
| 4.2 | Présentation de l'artérite à cellules géantes | 108 |
| 4.3 | Première Méthode : Modèle bayésien pour la conduite d'un essai de non-infériorité dans le cas des maladies rares | 110 |
| 4.3.1 | Introduction de la méthode | 110 |
| 4.3.2 | Modèle bayésien | 111 |
| 4.3.3 | Détermination d'une distribution <i>a priori</i> sur la base d'opinion d'experts. | 113 |
| 4.3.4 | Force de l'avis <i>a priori</i> des experts | 117 |
| 4.3.5 | Implémentation numérique sur le logiciel R | 119 |
| 4.3.6 | Loi <i>a posteriori</i> pour un échantillon fixé | 121 |
| 4.3.7 | Probabilité de non-infériorité en incorporant des données d'études historiques | 125 |
| 4.4 | Deuxième méthode : Méthodes de différences de deux lois bêta | 128 |
| 4.4.1 | Principe de la méthode | 128 |
| 4.4.2 | Choix des <i>a priori</i> en se basant sur P1 et P2 | 128 |
| 4.4.3 | Calcul de la loi <i>a posteriori</i> | 131 |
| 4.4.4 | Incorporation des données d'un essai voisin et résultats | 132 |
| 4.4.5 | Interprétation | 134 |
| 4.5 | Troisième méthode : Méthodes ϵ -contamination | 135 |
| 4.5.1 | Principe de la méthode | 135 |
| 4.5.2 | Risque bayésien | 137 |

| | | |
|-------|---|-----|
| 4.5.3 | Cas où il y a plusieurs informations de mesure différentes provenant d'autres études | 138 |
| 4.5.4 | Présentation de la pseudo-polyarthrite rhizomélique et sa relation avec l'artérite à cellules géantes | 140 |
| 4.5.5 | Calcul des lois <i>a posteriori</i> avec la méthode ϵ -contamination | 142 |
| 4.5.6 | Calcul de la différence entre les traitements | 144 |
| 4.5.7 | Interprétation | 145 |
| 4.6 | Annexe du chapitre 1 | 151 |
| 4.7 | Annexe du chapitre 2 | 167 |
| 4.7.1 | Interchangeabilité et mise en place des modèles hiérarchiques | 167 |
| 4.7.2 | Preuve du Théorème 2.61 | 169 |
| 4.8 | Chapitre 4 | 171 |
| 4.9 | Algorithme de calcul | 171 |
| | Bibliographie..... | 177 |

Chapitre 1

Les méthodes de gestion des maladies rares et des essais cliniques sur petits échantillons

1.1 Introduction

Les maladies rares constituent un ensemble hétérogène de conditions qui affectent différents organes qui varient selon un gradient de rareté.

Il existe plusieurs définitions de maladies "rares" ou "orphelines" et qui peuvent varier selon les pays. Les autorités médicales dans le monde s'appuient généralement sur la faible prévalence d'une maladie et sur le fait qu'il existe peu de traitements spécifiques à ces maladies. Une maladie est dite rare en Europe lorsqu'elle affecte moins de 1 personne sur 2000, tandis qu'aux États-Unis elle est définie lorsqu'elle affecte moins de 1 Américain sur 200.000 [<http://www.eurordis.org>].

En revanche, la prévalence peut varier selon le pays, de ce fait on peut avoir des maladies qui sont rares dans une région ou pays et fréquentes dans un autre. Ainsi, la drépanocytose, une maladie génétique touchant l'hémoglobine, est fréquente en Afrique et en Inde et rare en France, soit une prévalence en 2007 d'un nouveau-né atteint pour 2065 naissances. Il existe également des maladies fréquentes qui présentent des variantes rares. Ainsi, le diabète néonatal est une maladie fréquente, néanmoins il présente des variantes rares comme certains diabètes génétiques de l'enfant.

Bien que presque toutes les maladies génétiques soient des maladies rares, les maladies

rare ne sont pas toutes des maladies génétiques. Ainsi, 20% des maladies rares sont le résultat d'infections, d'allergies ou d'autres origines, ou bien sont des maladies dégénératives.

Les signes d'apparition de plusieurs maladies rares peuvent être détectés avant ou dès la naissance ou bien durant l'enfance, comme c'est le cas de la neurofibromatose, de la maladie de Krabbe ou de la maladie de Batten. Selon le ministère de la Santé, 35% des décès surviennent avant l'âge de un an, 10 % des décès entre 1 et 5 ans et 12 % entre 5 et 15 ans. Cependant, 50% des maladies rares apparaissent durant l'adolescence ou à l'âge adulte, comme la maladie de Huntington, une maladie neurologique rare.

De nos jours, on estime qu'il existe entre 6 000 et 7 000 maladies rares distinctes et des nouvelles maladies rares sont régulièrement décrites dans la littérature médicale. Le nombre des maladies rares dépend du degré de spécificité utilisé en classifiant les différents troubles. Chacune des maladies trouvées est caractérisée par une grande diversité de troubles et de symptômes qui varient non seulement d'une maladie à l'autre, mais aussi d'un patient à l'autre souffrant de la même maladie. De nombreux diagnostics révèlent une grande diversité de sous-catégories pour une même maladie. Les maladies rares diffèrent également dans leur gravité. La plupart mettent en danger la vie des malades, tandis que d'autres sont compatibles avec une vie normale si elles sont diagnostiquées à temps et traitées correctement [<http://www.eurordis.org>].

La notion "rare" donnée à chacune de ces maladies n'empêche pas qu'on estime à plus de 30 millions de personnes atteintes de maladies rares en Europe. Dans l'UE, les patients atteints de maladies rares représentent entre 6 et 8% de la population.

Le domaine des maladies rares souffre de déficit médical et de connaissances scientifiques, puisque la véritable recherche et la politique de santé publique concernant les questions liées à ce domaine ne sont apparues que récemment. De nombreux obstacles existent pour faire progresser les connaissances et les options de traitement pour les maladies rares. Le défi le plus évident pour mener une recherche rigoureuse est le petit nombre de participants admissibles pour une étude donnée. En outre, la dispersion géographique des patients, le manque de connaissance sur l'évolution clinique de la maladie, et le manque de traitement de comparaison approprié entravent la production de preuves. De plus, peu d'informations sont connues sur l'évolution clinique de nombreuses maladies rares et peu d'options thérapeutiques existent.

Au cours des dernières années, des méthodes innovantes d'enquêtes épidémiologiques et d'essais cliniques ont été développées; et elles sont prometteuses pour promouvoir une recherche plus efficace.

Pour cette raison, nous avons systématiquement recherché dans la littérature les publi-

cations concernant les caractéristiques des différentes méthodes qui ont été utilisées pour l'étude des maladies rares. L'objectif est d'identifier des approches novatrices pour la recherche qui ont été, ou peuvent être, utilisées afin de surmonter les difficultés méthodologiques inhérentes à l'étude des maladies rares.

Cette première recherche nous permettra d'avoir une vue exhaustive sur les méthodes statistiques classiques utilisées, avec leurs avantages et leurs inconvénients.

1.2 Les méthodes non bayésiennes utilisées dans la littérature

1.2.1 Méthodes de recherche

L'objectif de cette recherche est d'identifier les approches méthodologiques utilisées dans le cas de l'étude des maladies rares.

Le mécanisme de recherche est fondé sur quatre phases.

1.2.2 Choix des mots clés

Il est important de choisir soigneusement les mots clés pour préciser la recherche. Nous avons effectué une recherche de mots clés associés aux études de maladies rares en utilisant plusieurs bases de données et via la recherche sur les sites associés aux maladies rares (exemple : Eurordis, Orphanet). La liste originale des mots utilisés dans notre revue de la littérature sont : "Rare diseases", "Clinical trials", "Statistical methods", "Small sample", "Epidemiology", "Research design".

1.2.3 Recherche dans les bases de données.

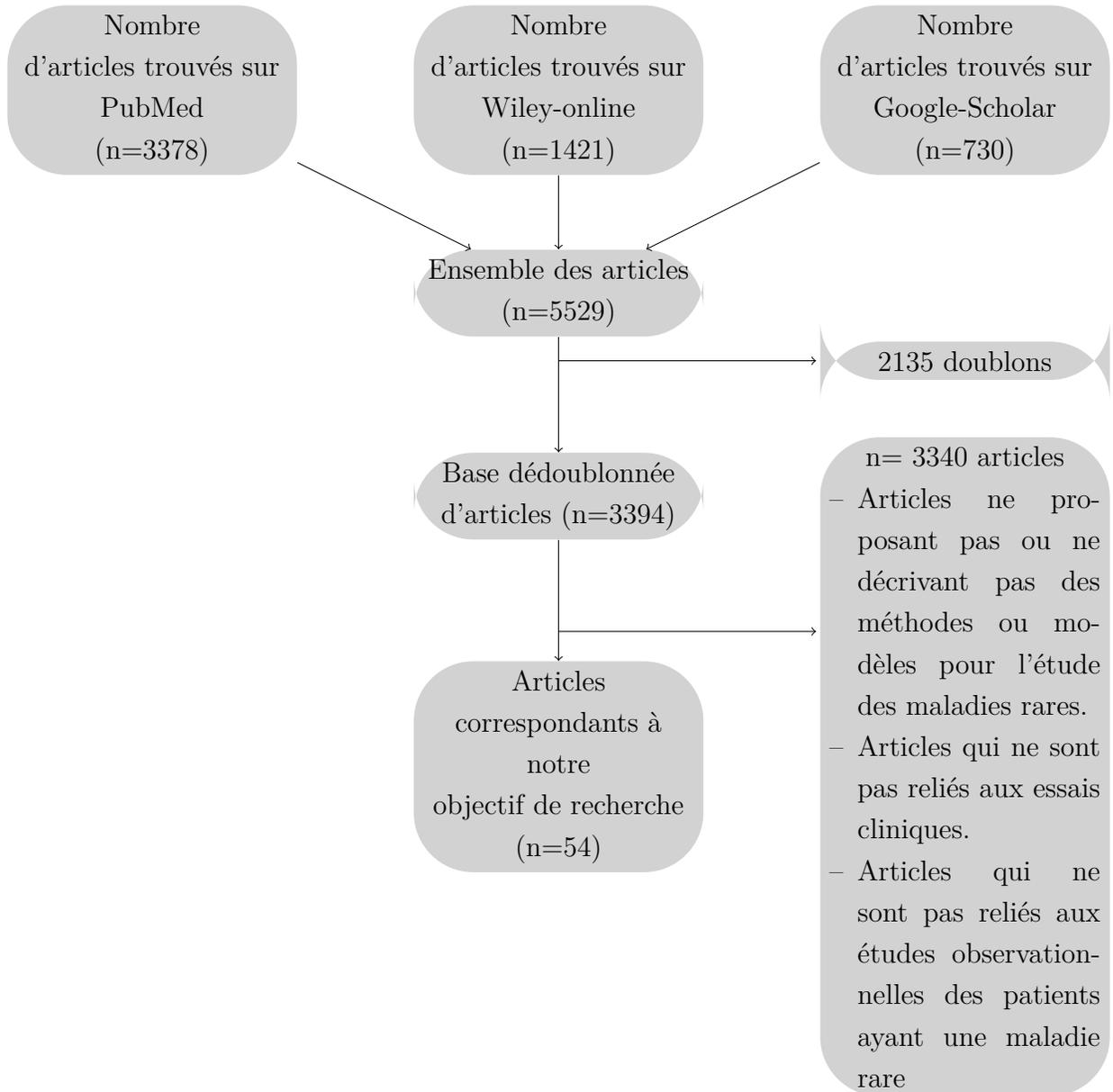
La phase 2 est la recherche elle-même. Nous avons utilisé en premier lieu les mots clés prédéfinis pour interroger trois bases de données : PubMed, Google Scholar et Wiley online Library. La recherche dans ces bases était limitée aux articles en anglais, mais sans limitation sur des années spécifiques. Ensuite, nous avons relancé une nouvelle recherche en utilisant les mots des méthodes ou modèles comme : "cross over design", "adaptive randomisation", "Bayesian Model" (qui seront détaillés dans le chapitre suivant).

La dernière étape de recherche a consisté à analyser toutes références additionnelles dans la bibliographie ou liste de références dans les articles trouvés.

1.2.4 Revues de la littérature

Nous avons utilisé "Zotero" (<https://www.zotero.org/>) pour noter tous les articles pertinents trouvés dans la totalité des bases (5529 articles) puis nous avons supprimé les doublons (il nous reste 3394 articles). Les articles restants ont été examinés attentivement, le premier pas était le criblage préliminaire et la revue des titres et des résumés pour exclure les articles trop généraux et sans rapport avec notre objectif, après que le criblage préliminaire et la revue ont été complétés, le reste des publications a été analysé, par la

suite nous avons supprimé les articles qui sont spécifiques à un cas individuel de maladies rares et ceux qui ne proposent ou ne décrivent pas des méthodes pour les études des maladies rares. Finalement, nous avons supprimé les articles qui n'étaient pas reliés à des essais cliniques ou aux études observationnelles de patients atteints de maladies rares.



1.2.5 Extraction des articles

Nous avons finalement retenu 54 articles concernant les modèles et méthodes utilisées dans l'étude thérapeutique sur les maladies rares. La plupart ont été publiés entre 1995 et 2015 et ces articles ont été sauvegardés dans JabRef, un logiciel de gestion bibliographique libre.

On identifie à partir des articles retenus, 21 stratégies pour étudier les maladies rares. Les méthodes innovantes trouvées dans les essais cliniques minimisent la taille de l'échantillon requise et maximisent la proportion des patients qui reçoivent le traitement actif.

Ces approches et méthodes ont été classées en trois groupes : les conceptions d'étude, les modèles adaptatifs et les stratégies statistiques.

1.3 Les conceptions d'étude

1.3.1 Modèle de démarrage retardé

Le concept de l'essai d'un modèle de démarrage retardé a été développé dans le domaine de la maladie d'Alzheimer, principalement par Paul Leber de la Food and Drug Administration des États-Unis [Clarke, 2008].

Le modèle de démarrage retardé est un design dans lequel les patients sont randomisés pour recevoir le même traitement actif, mais en commençant à des moments différents, ce qui entraîne deux phases de traitement : la phase *I* qui est une période contrôlée par un placebo suivie par la phase *II* qui est la période de démarrage retardé. Au cours de la phase *I*, les patients reçoivent soit un traitement actif soit un placebo, durant la phase *II* les patients du le groupe contrôle qui ont reçu le placebo dans la phase *I* sont commutés pour recevoir un " démarrage retardé" du traitement actif et les patients qui ont reçu le traitement actif continuent à recevoir ce traitement et on les appelle "patients de démarrage précoce" [Liu-Seifert et al., 2015].

Planifier et concevoir un modèle de démarrage retardé nécessite une connaissance approfondie du processus de la maladie étudiée, une sélection et une implémentation très attentive des méthodes statistiques. Les données manquantes peuvent aussi constituer un problème important. Les participants du groupe contrôle sont plus susceptibles d'abandonner l'étude en raison de l'absence d'effet du traitement, conduisant à un taux d'abandon différent entre le groupe de traitement et le groupe de référence. Ce protocole doit préciser les méthodes statistiques appropriées qui peuvent contrôler ce problème. L'analyse des covariables et les scores de propension peuvent également être utilisés pour vérifier un

déséquilibre entre les groupes témoin et contrôle.

Dans ce modèle, la durée de la phase *I* devrait être suffisamment longue pour donner le temps nécessaire aux effets du traitement de se manifester sur les symptômes de la maladie. Idéalement, à la fin de cette période, toutes différences significatives entre les deux groupes (le cas échéant) reflètent les différences sur les effets sur les symptômes ou les effets sur l'évolution de la maladie, ou bien les deux. Les patients sont ensuite suivis pendant une période de temps prédéterminée. Tous les bénéfices des effets symptomatiques devraient être égaux dans les deux groupes, et toutes les différences restantes être attribuables à l'effet du traitement modificateur de la maladie [D'Agostino, 2009].

Pendant toute la durée de l'étude, cette randomisation au groupe de traitement est en double aveugle, c'est-à-dire que ni les patients ni les cliniciens ne connaissent la nature du traitement délivré.

Les modèles de démarrage retardé sont utiles quand il y a besoin d'une étude clinique appropriée qui mesure de façon fiable la progression de la maladie et qui ne soit pas confondue avec les effets de l'étude d'intervention sur les symptômes.

Pendant la phase *I*, une approche statistique est nécessaire pour montrer qu'il y a une différence en termes d'effets sur les symptômes, mais pas nécessairement une différence en termes d'effets modificateurs de la maladie [Liu-Seifert et al., 2015]. Toutefois, afin de tester uniquement les effets sur les symptômes, il serait suffisant que le traitement actif ait une meilleure performance que le placebo ou que le traitement de référence à la fin de la phase *I*. Cela pourrait être testé statistiquement en examinant les données seulement à la fin de la phase *I* ou bien en faisant la moyenne d'une partie ou de la totalité des données dans la phase *I*. La détermination de l'approche statistique doit être faite avant d'examiner les données. En outre, les méthodes utilisées pour l'analyse des données en phase *I* nécessitent une erreur alpha appropriée et ajustée pour des tests multiples. Idéalement, à la fin de la phase *I*, il faut sélectionner un test statistique sur la variable primaire qui établit que les effets du placebo ou du traitement de référence s'il existe et celui du traitement expérimental sont différents.

La phase *II* comprend deux hypothèses de recherche et au moins deux tests statistiques. Le premier, à la fin de l'étude (la fin de la phase *II*), un test statistique doit être effectué pour déterminer s'il y a une évolution moins rapide dans le groupe placebo que dans le groupe traitement actif, cela permettrait d'établir l'hypothèse de recherche que le bénéfice observé à la fin de la phase *I* persiste.

Le deuxième test statistique doit être effectué pour établir l'hypothèse que les effets modificateurs de la maladie persistent et ne sont pas diminués au cours de la période de l'étude dans le groupe traitement expérimental versus le groupe placebo ou du traitement

de référence [D'Agostino, 2009].

Dans la littérature très peu de données ont été publiées sur les études de démarrage retardé.

Deux études ont été menées pour évaluer les effets de l'initiation précoce et tardive de la rasagiline sur l'évolution de la maladie de Parkinson [Parkinson, 2004], [Olanow et al., 2009]. Ces études ont suggéré que les effets de la rasagiline ne peuvent pas être pleinement expliqués par son effet symptomatique et pourraient être dus à une activité propre du médicament qui modifierait l'expression de la maladie et que des études de longue durée sont nécessaires.

Enfin, il n'y a pas eu de mise en oeuvre de méthodes quantitatives pour déterminer si les différences entre les patients de démarrage précoce versus retardé étaient dues au hasard.

1.3.2 Les essais croisés

Les essais croisés sont des essais randomisés dans lesquelles les patients sont exposés à des séquences de différents traitements. Le but est de trouver les différences entre ces traitements (comme pour un RCT normal). Dans les études croisées, chaque patient est exposé aux deux traitements dans différents intervalles de temps et ils sont randomisés dans l'ordre par lequel ils reçoivent le traitement expérimental ou le traitement contrôle (ou le traitement de référence). Dans ces études la comparaison est réalisée en tenant compte des traitements et de leurs périodes d'allocation donc on n'a pas à s'inquiéter des déséquilibres des facteurs pronostiques [Maclure, 1991], [Brophy et al., 2003], [Lagakos, 2003], [Griggs et al., 2009], [Mann, 2003]. Les essais croisés sont particulièrement bien adaptés aux études des traitements pour les maladies chroniques dont les traitements apportent un soulagement immédiat des symptômes.

Mais ces essais ne peuvent généralement pas être utilisés pour étudier les traitements qui ont des effets curatifs ou des conditions qui évoluent rapidement.

Beaucoup de maladies rares sont des maladies chroniques qui progressent au cours du temps. Les variations de la maladie au fil du temps peuvent causer des biais dans les essais croisés [Lagakos, 2003].

Les patients recevant les traitements peuvent exprimer leurs préférences pour ou contre les traitements. Comme chaque patient reçoit les deux ou plusieurs interventions, les essais croisés ne demandent en général pas plus de $\frac{N}{2}(1-\rho)$ patients pour avoir la même puissance statistique qu'un essai à groupes parallèles, avec $\rho =$ corrélation entre la variable principale, $N =$ nombre total de patients pour un essai en groupes parallèles [Gupta et al., 2011]

[Maclure, 1991].

De plus comme un même patient reçoit tous les traitements séquentiellement, la variation dans les réponses pour un seul patient est moindre que celle qui existe entre plusieurs patients (variable intra patient et inter patient).

Quand la variation intra patient est plus petite que la variation inter patient, on dit qu'il y a corrélation entre les réponses des patients aux différents traitements [Gupta et al., 2011].

Pour exploiter cette corrélation, les essais croisés doivent être analysés en utilisant une méthode d'analyse spécifique pour les données du même patient à différent point du temps (test de Student ou de McNemar test).

De plus l'analyse peut présenter d'effet période et l'effet ordre ("carryover") [Wellek and Blettner, 2012]

Dans les plus part des plans croisés de base impliquant deux traitements, il y a une possibilité d'avoir un effet "carryover" d'une phase de traitement à l'autre. Un effet "carryover" intervient lorsque le traitement donné dans la première période a un effet qui persiste jusqu'à la seconde période. Par exemple, dans un essai contrôlé contre placebo, les patients prenant le placebo en premier puis prenant le traitement actif peuvent ne pas être affectés par l'effet "carryover". Mais les patients recevant en premier le traitement actif peuvent présenter un effet rémanent durant la partie initiale de la deuxième période.

L'effet "carryover" réduit l'efficacité apparente du traitement pour les patients dans le deuxième groupe, et en conséquence pour tous les patients et périodes combinés.

Une période "washout" dans laquelle aucun traitement n'est donné entre les périodes du traitement peut réduire ou annuler l'effet du risque de "carryover". La période de "washout" doit être suffisamment longue pour s'assurer que l'effet rémanent du premier traitement n'influence pas l'effet du second traitement donné dans la phase suivante, mais cela va encore prolonger le temps de l'étude [Senn, 2006].

Lorsque les chercheurs ne peuvent pas estimer avec précision la durée dont on a besoin pour la période "washout", des modèles plus complexes sont nécessaires impliquant plus de deux périodes (par exemple : ABB/ BAA) ou plus de deux traitements (par exemple : AA, BB, AB, BA) [Nason and Follmann, 2010], [Gupta et al., 2011].

Une méthode recommandée de l'analyse des essais croisés est de tester l'effet "carryover" ; s'il est significatif, alors on rejette les données de la deuxième période, et on analyse les données de la première période comme si on était dans un essai en groupes parallèles.

Bien qu'il soit fortement souhaitable que les auteurs rapportent les résultats de chaque traitement séparément, [Freeman, 1989] a montré que chaque stratégie expose à une réponse biaisée.

De plus, [Senn, 2006] ont montré que l'utilisation du modèle d'essais croisés est effective-

ment construite sur l'hypothèse qu'il y a un "carryover" minimal de l'effet de traitement dans la période suivante.

La planification d'un essai croisé devrait inclure un calcul bien fondé des tailles d'échantillons. Dans le cas des études croisées, ceci est le test de différence entre les effets du traitement. La planification de l'essai sera généralement faite sous l'hypothèse que la phase de "washout" est suffisamment longue pour exclure l'effet de "carryover". En principe, la procédure nécessaire pour le calcul de la puissance et de la taille des échantillons pour un essai croisé est le même que celui qui est familier du t-test non apparié. La seule différence réside dans la spécification des hypothèses dans lesquelles la puissance prédéfinie (par exemple 80%) devrait être obtenue [Wellek and Blettner, 2012]. Une question importante est de savoir si le design croisé a une efficacité supérieure ou inférieure par rapport à une étude normale à deux bras fournissant des données d'une période d'étude unique. L'efficacité se réfère ici à la taille des échantillons requis par les deux modèles pour atteindre la même puissance dans des conditions identiques.

1.3.2.1 Les étapes de l'analyse statistique de confirmation d'un essai croisé

Soit C_{1i} et C_{2i} les résultats d'un patient i de la séquence A-B dans la période 1 ou 2 respectivement,

Et D_{1j} et D_{2j} les résultats d'un patient j de la séquence B-A dans la période 1 ou 2 respectivement.

Les sommes intra sujet des résultats des deux périodes sont présentés respectivement par

$$X_i(C) = C_{1i} + C_{2i} \text{ et } X_j(D) = D_{1j} + D_{2j}.$$

En addition, on présente les différences intra sujet des résultats de la période 1 et la période 2 par

$$Y_i(C) = C_{1i} - C_{2i} \text{ et } Y_j(D) = D_{1j} - D_{2j};$$

On note m et n le nombre de patients dans les séquences A-B et B-A respectivement, avec $N = m + n$.

Tout d'abord, il faut faire un pré-test pour vérifier si l'hypothèse que l'effet de "carryover" est négligeable. Ce pré-test doit être exécuté comme un t-test "ordinaire" non apparié avec $X_1(C), \dots, X_m(C)$ et $X_1(D), \dots, X_n(D)$ comme les deux échantillons.

La statistique du test est alors calculée selon l'équation suivante :

$$T = \sqrt{\frac{mn}{N}} \frac{\bar{X}(C) - \bar{X}(D)}{\sqrt{(S_{XC} + S_{XD})/(N - 2)}} \quad (1.1)$$

avec :

$$\begin{aligned} \bar{X}(C) &= (X_1(C) + \dots + X_m(C))/m, \\ \bar{X}(D) &= (X_1(D) + \dots + X_n(D))/n, \\ S_{XC} &= (X_1(C) - \bar{X}(C))^2 + \dots + (X_m(C) - \bar{X}(C))^2, \\ S_{XD} &= (X_1(D) - \bar{X}(D))^2 + \dots + (X_m(D) - \bar{X}(D))^2. \end{aligned}$$

En suite, il faut tester les différences entre les effets du traitement. Pour cela, on suit exactement le schéma défini pour le pré-test (équation (1.1)).

$$T = \sqrt{\frac{mn}{N}} \frac{\bar{Y}(C) - \bar{Y}(D)}{\sqrt{(S_{YC} + S_{YD})/(N - 2)}}$$

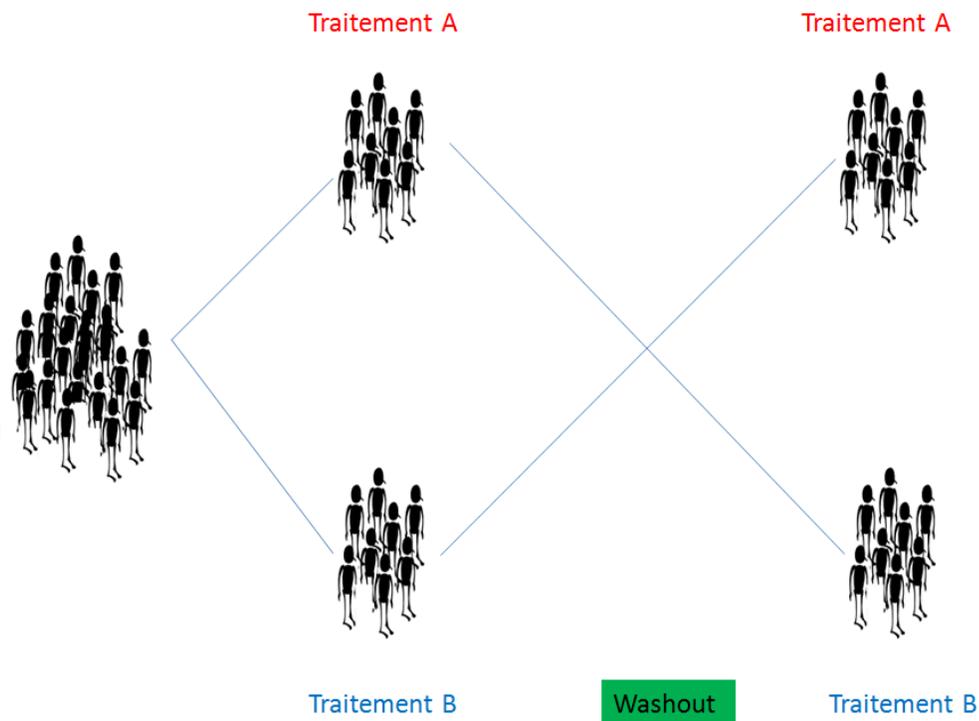


Figure 1.1 – Schéma représentatif des essais croisés

1.3.3 Les essais de type N-de-1

Une étude N-de-1 est un type spécial de l'essai croisé où l'intervention est évaluée sur un seul patient. Ces essais se composent de périodes de traitements expérimentaux/contrôle répétés un certain nombre de fois. L'ordre de traitement est attribué de façon aléatoire à l'intérieur de chaque période de traitement paire. Ce design est connu comme un essai randomisé structuré multi croisé d'un seul patient [Lillie et al., 2011]. Ces essais croisés à un patient ont été présentés pour aider à la prise de décisions et pour se concentrer à la réponse individuelle aux traitements.

Alors que l'inférence statistique ne peut être faite sur la base d'un seul essai N-de-1, les résultats de plusieurs de ces études peuvent être regroupés dans une série de cas ou même une méta-analyse quantitative ([Berlin, 2010]). La plupart des études publiées sur les essais N-de-1 résumet et comparent les résultats entre les essais provenant d'une série conçue d'une manière prospective. Pour combiner les recherches sur les essais N-de-1, on pourrait utiliser les résultats, s'ils sont disponibles, des essais déjà réalisés indépendamment pour la gestion des traitements personnels. Au lieu de la première, on peut entreprendre des séries prospectives avec l'objectif de la recherche d'une population claire. Ou bien on pourrait utiliser une combinaison de sources d'essais.

Chaque essai N-de-1 fournit des informations non seulement sur les comparaisons des traitements intra patients, mais aussi sur la variation des réponses de chaque traitement au sein de chaque patient.

Les informations collectées de chaque essai peuvent être combinées dans une méta-analyse, ou bien on a recours aux modèles linéaires mixtes en utilisant le maximum de vraisemblance, ou bien on utilise les modèles bayésiens hiérarchiques avec ou sans distributions *a priori* informatives [Zucker et al., 2010].

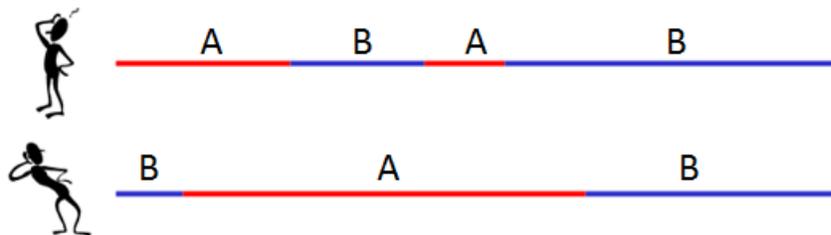


Figure 1.2 – Schéma représentatif des essais de type N-de-1 avec les traitements A et B

1.3.4 Plan factoriel

Les chercheurs qui étudient les maladies rares ont essayé de faire face au petit groupe de participants expérimentaux. Certains ont proposé ou ont fait des modifications des méthodes utilisées dans les essais traditionnels. Par exemple, lorsque l'on considère la période de traitement, le choix d'une durée plus longue de l'essai peut réduire les besoins en taille de l'échantillon en récoltant plus d'événements parmi les participants à l'essai (cas de la survie).

Pour cette raison certains chercheurs ont tenté de réduire la taille de l'échantillon en abordant de multiples options de traitements dans une étude factorielle, dans lequel deux (ou plus) des comparaisons de traitement sont effectuées simultanément [Cornu et al., 2013], [Nony et al., 2014], [Griggs et al., 2009].

Les études factorielles sont similaires aux modèles des études croisés, mais à la différence prête que les modèles factoriels fournissent des réponses à plusieurs questions dans la même population d'étude et permettent aux chercheurs d'évaluer plus d'une intervention en une seule expérience.

Cela réduit le nombre total de patients requis pour répondre à toutes les questions, mais ne réduit pas le nombre de patients requis pour répondre à chaque question.

Ce type d'étude suppose également qu'il n'y a pas d'interaction entre les (deux) traitements. Nous entendons par interaction que l'effet d'un traitement A sur un groupe de comparaison est la même indépendamment du fait que le patient a reçu un traitement B ou non.

Ce modèle factoriel permet de mesurer un effet ou une interaction qui autrement n'aurait pas été apparente.

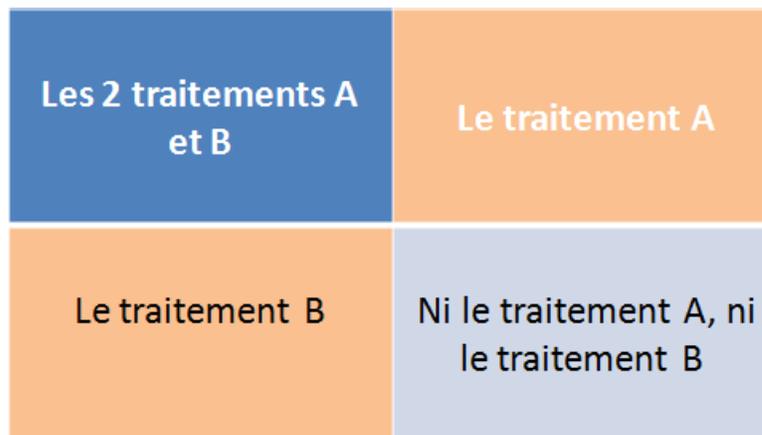


Figure 1.3 – Schéma représentatif du plan factoriel

1.3.5 Design de retrait aléatoire et modèle d'échappement précoce

Ces modèles visent à appliquer des «stratégies d'enrichissement» à la phase de conception de l'étude, afin de sélectionner un sous-groupe de patients dans lequel la détection d'un effet thérapeutique est plus probable que dans toute la population des malades.

Le design de retrait aléatoire est l'une des approches les plus connues. Ce modèle a deux phases [Katz, 2009] :

La première phase est utilisée pour identifier les patients qui ont répondu au traitement en inscrivant les participants à l'étude dans un essai ouvert où tous les patients reçoivent le traitement étudié. Une fois que les patients qui ont répondu au traitement sont identifiés, ils sont inscrits dans la deuxième phase de retrait aléatoire, au cours de laquelle ils sont répartis au hasard soit pour continuer à recevoir le traitement expérimental, soit passer à un traitement de contrôle (placebo). Les patients des deux groupes sont suivis régulièrement sur la base de critères cliniques ou biologiques jusqu'à ce que l'effet sur une variable intermédiaire apparaisse.

L'analyse de l'étude est effectuée en utilisant uniquement les données de la deuxième phase, ce qui inclut uniquement les sujets qui sont "susceptibles" de répondre à la thérapie [Abrahamyan et al., 2016].

Le design du retrait aléatoire permet souvent aux sujets qui ont été retirés de la thérapie (sujets du groupe témoin) d'être remis sur ce qui a été une thérapie expérimentale efficace, après avoir atteint le point final de l'étude.

Cette conception est largement utilisée dans l'évaluation de thérapies pour la douleur chronique, les affections psychiatriques et dans d'autres populations hétérogènes où l'on s'attend à ce que le traitement ne soit visible que dans un sous-ensemble de patients [Straube et al., 2008] .

Les défis méthodologiques avec cette conception comprennent l'établissement de la durée correcte pour la première phase (et si l'augmentation de la dose doit avoir lieu au cours de cette période), une identification précise des «intervenants» pendant cette phase (c'est-à-dire déterminer si une réponse représente réellement une rémission ou une amélioration dans l'activité de la maladie), et des méthodes d'imputation des données pour les valeurs manquantes pour les retraits des événements indésirables à long terme [Katz, 2009] [Abrahamyan et al., 2016]. Les patients qui sont ensuite randomisés dans le groupe contrôle peuvent subir des effets "carryover".

Enfin, ce design limite la généralisation pour les patients qui présentent des caractéristiques différentes systématiquement de ceux identifiés dans la première phase (c'est-à-dire

des «répondeurs»). En conséquence, une attention particulière devrait être accordée à la validité externe lorsqu'on discute des implications des résultats sur la pratique.

Nous présentons le modèle d'échappement précoce qui est une autre façon de minimiser la durée d'exposition d'un individu à un placebo et qui est utilisé couramment dans les maladies rares et/ou les petits essais.

Cette méthode permet aux patients d'arrêter ou d'échapper aux essais randomisés attribués lorsque le patient ne répond pas aux paramètres prédéfinis.

Une des difficultés importantes de ces essais est de définir un critère binaire d'évaluation échec/succès, ainsi que la perte de puissance rapide de ces études en cas où on a un nombre important de perdus de vue [M.D, 1994].

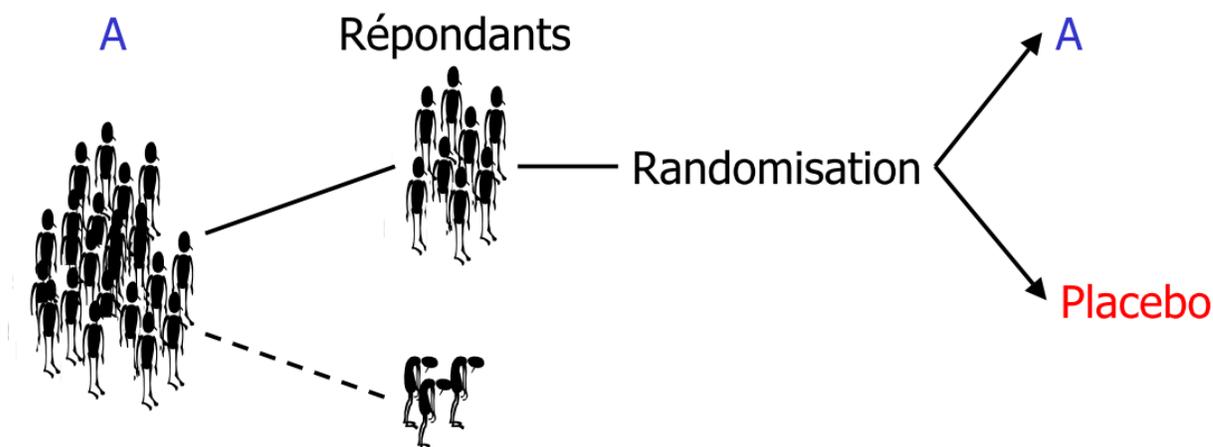


Figure 1.4 – Schéma représentatif d'un design de retrait aléatoire

1.3.6 Conception randomisée de la phase placebo

La conception randomisée de la phase placebo (*CRPP*) a été développée dans le but principal d'améliorer l'acceptabilité d'entrer dans un essai en diminuant efficacement la durée pendant laquelle chaque sujet est exposé au placebo [Feldman et al., 2001]. Il a été conçu pour des traitements susceptibles de provoquer une rémission ou une réponse durable (et ne seraient donc pas adaptés aux essais croisés).

Comme dans l'essai randomisé contrôlé (*ECR*) des groupes parallèles, les patients du "*CRPP*" sont d'abord répartis au hasard soit dans un groupe expérimental, soit dans un groupe témoin. Cependant, après une courte période de temps fixe (appelée phase placebo), tous les patients du groupe témoin passent au traitement expérimental. Ce design repose sur l'hypothèse que si le traitement est efficace, les patients qui l'ont reçu

plus tôt (c'est-à-dire ceux qui ont été initialement alloués au groupe expérimental) répondront en moyenne plus tôt. À la fin de l'essai, le temps de réponse moyen est comparé entre les deux groupes, le plus souvent en utilisant des méthodes d'analyse de survie [Abrahamyan et al., 2016].

Une étude de simulation évaluant la performance du RPPD contre le "ERC" du groupe parallèle dans différentes distributions de la durée de la réponse [Abrahamyan et al., 2011] a révélé que sous des distributions spécifiques (par exemple, lognormal, Weibull), la "CRPP" peut être très proche en efficacité (puissance statistique) à l'"ERC" des groupes parallèles. Une autre étude de simulation a établi une surveillance interne efficace et des règles d'arrêt précoce pour la "CRPP" [Shook, 2010]. En conséquence, si la "CRPP" réduit les obstacles au recrutement, elle peut offrir un fort avantage dans la pratique avec un inconvénient statistique limité.

Une considération du design spécifique est l'établissement de la durée optimale de la phase placebo et de la durée d'essai globale. Une durée plus longue de la phase placebo diminuera la taille de l'échantillon requise, mais augmentera les chances d'abandon des patients. La durée optimale, cependant, ne doit être définie qu'après avoir soigneusement pris en compte les préférences du patient, la progression de la maladie, la taille minimale d'effet important cliniquement et la répartition statistique du temps de réponse au traitement [Abrahamyan et al., 2011].

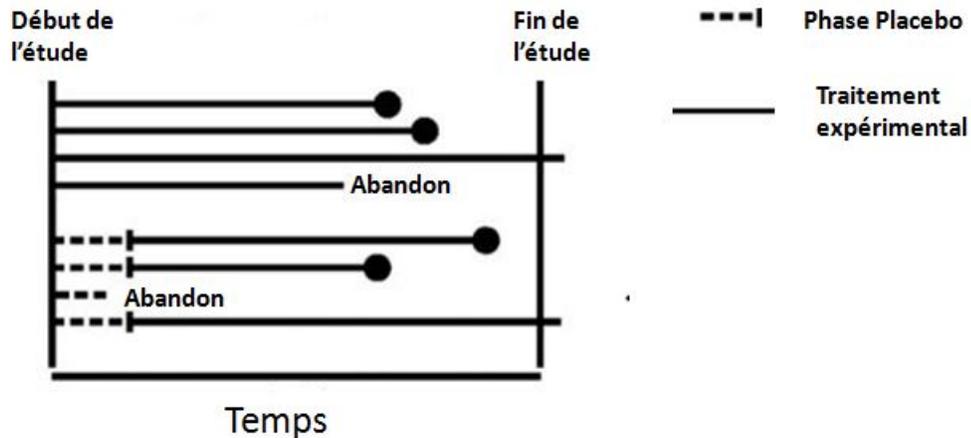


Figure 1.5 – Schéma représentatif de la conception randomisée de la phase placebo

1.4 Modèles adaptatifs

Malgré le désir incontestable d'obtenir des réponses liées à n'importe quelle intervention d'un essai randomisé contrôlé (ERC), il n'est parfois pas faisable de procéder à un ERC assez grand, comme dans le cas des maladies rares où la disponibilité d'un petit nombre de patients en raison de la rareté de la maladie rend difficile de mener un ERC. Dans ce cas on peut utiliser les modèles de randomisation adaptative qui permettent des modifications apportées à l'essai et/ou aux procédures statistiques des essais cliniques en cours. Dans les dernières années, l'utilisation des modèles adaptatifs dans les essais cliniques est devenue très populaire en raison de flexibilité et efficacité. Les modèles adaptatifs sont définis comme un modèle qui permet les modifications à l'essai et/ou aux procédures statistiques de l'essai après son initiation sans influencer sa validité et son intégrité. Compte tenu du niveau de flexibilité, ces conceptions d'essais sont également appelées comme "conceptions flexibles" [Mahajan and Gupta, 2010].

Les conceptions d'essais peuvent être modifiées en fonction de l'accumulation des données de l'essai, ou même des données extérieures. Dans l'utilisation de schémas adaptatifs, un essai peut être arrêté à mi-chemin, s'il y a des indications suffisantes de manque de sécurité ou d'efficacité. De même, la taille de l'échantillon peut être ré estimée, les bras de traitement inefficace abandonnés, de nouveaux bras ajoutés et les hypothèses de l'essai réévaluées. La mise en oeuvre des essais adaptatifs peut être difficile en raison de la complexité impliquée dans la planification de ces essais et de la validité statistique de ces adaptations.

1.4.1 Types de modèles adaptatifs

Différentes adaptations dans ces essais cliniques adaptatifs peuvent être groupées en 2 catégories : Adaptations dans les procédures des essais et adaptation dans les processus statistiques. Les adaptations dans les procédures des essais incluent les changements des critères d'éligibilité, de dosage, de durée du traitement, de paramètres de l'étude, de procédures de diagnostic, de critères d'évaluations, et d'évaluation des réponses cliniques. L'adaptation dans les processus statistiques peut inclure les changements dans la randomisation, les hypothèses de l'étude, la réalisation d'une analyse intermédiaire, et le plan d'analyse statistique.

Les designs utilisés dans le cas des maladies rares sont les suivants : modèles de randomisation adaptative (randomisation adaptative aux réponses, randomisation adaptative aux covariables), méthode de groupes séquentiels, un design de réestimation de la taille

d'échantillon, abandonner le perdant, un design adaptatif de recherche de dose, un design de biomarqueur adaptatif, un modèle adaptatif de commutation de traitement, une conception d'hypothèse adaptative (dans la table 4.5 on présente en résumé les différents types de modèles adaptatifs, ainsi que les problèmes et les défis associés).

Le plus grand avantage d'un essai adaptatif est qu'il permet des changements qui peuvent diminuer le temps de période d'essai. Diminuer la durée de la période d'essai entraîne des économies de coûts et de temps importantes. Un essai peut être arrêté tôt si les données accumulées suggèrent que le médicament n'est pas efficace ou présente un problème de sécurité majeur. Un essai peut également être arrêté si les données accumulées suggèrent que, en fonction du nombre de patients recrutés, l'efficacité du médicament peut être démontrée statistiquement. Dans d'autres situations, certains bras inefficaces peuvent être abandonnés tôt, la meilleure dose peut être identifiée au début de l'essai. Un autre avantage des modèles d'adaptation est que l'information même en dehors de l'essai peut être utilisée dans l'analyse intermédiaire, ce qui peut avoir un impact sur le procès.

Bien que les essais cliniques adaptatifs présentent de nombreux avantages, leur mise en oeuvre peut être difficile en raison de la complexité impliquée dans la planification de tels essais, les problèmes de logistique impliqués et la validité statistique des adaptations.

1.4.2 Les modèles de randomisation adaptative

Dans les modèles de randomisation adaptative, des modifications dans le plan de randomisation sont autorisées suivant les probabilités variées ou inégales des traitements affectés. Ces modèles sont couramment appliqués et comprennent le traitement adaptatif de randomisation, randomisation covariable adaptative et la randomisation de réponse adaptative. Le but de la randomisation adaptative peut être de parvenir à un design plus équilibré ou à réduire l'écart par rapport au ratio d'attribution du traitement cible en utilisant une probabilité d'attribution variée (randomisation adaptative de traitement) ou de minimiser déséquilibre dans les covariables initiales entre les groupes de traitement (randomisation Covariable adaptative) ou pour augmenter la proportion de patients affectés au traitement apparemment plus efficace tout en réduisant les effectifs de l'essai (randomisation réponse adaptative). Nous présentons dans les sections suivantes la formulation mathématique des essais randomisés [Hu and Rosenberger, 2006].

1.4.3 Formulation mathématique des essais randomisés

Considérons un essai de n patients, dont chacun doit recevoir au hasard l'un des traitements K . Une séquence de randomisation est une matrice $T = (T_1, \dots, T_n)'$, tel que

$$\begin{cases} T_i = e_j \text{ avec } i = 1, \dots, n \text{ et } j = 1, \dots, K \\ \text{et } e_j \text{ est un vecteur nul avec un 1 sur la position } j \end{cases}$$

Donc,

$$T = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}$$

Nous allons présenter les propriétés importantes de la séquence de randomisation, qui implique habituellement de dériver des propriétés asymptotiques des proportions d'allocations présentées par $\frac{N(n)}{n}$, où $N(n)$ = nombre de réponses pour chaque traitement, d'où :

$$N(n) = (N_1(n), \dots, N_K(n)) \text{ et } N_j(n) = \sum_{i=1}^n T_{ij}.$$

Soit, une matrice de variables de réponses $X = (X_1, \dots, X_n)'$, où $X_i = (X_{i1}, \dots, X_{iK})$ représentent la séquence des réponses observées si chaque traitement a été assigné au $i^{\text{ème}}$ patient indépendamment [Hu and Rosenberger, 2006].

Nous supposons maintenant $\mathcal{F}_n = \mathcal{T}_n \otimes \mathcal{X}_n \otimes \mathcal{Z}_n$, avec :

$$\begin{cases} \mathcal{T}_n = \sigma \{T_1, \dots, T_n\} \text{ est une sigma-algèbre générée par les premières} \\ \text{affectations du traitement} \\ \mathcal{X}_n = \sigma \{X_1, \dots, X_n\} \text{ est une sigma-algèbre générée par les premières} \\ \text{réponses du traitement} \\ \mathcal{Z}_n = \sigma \{Z_1, \dots, Z_n\} \text{ est une sigma-algèbre générée par les premières} \\ \text{covariables du traitement} \end{cases}$$

Une procédure de randomisation est définie par :

$$\Phi_n = E(T_n | \mathcal{F}_{n-1}) \text{ où } \Phi_n \text{ est } \mathcal{F}_n \text{ mesurable}$$

Nous pouvons décrire Φ comme la probabilité conditionnelle d'attribuer les traitements K au $n^{\text{ème}}$ patient, conditionnellement aux attributions, aux réponses et aux vecteurs de covariables du patient précédent $n-1$, et le vecteur de covariable du patient actuel n .

Nous allons décrire dans ce qui suit cinq types de procédures de randomisation : la randomisation complète, la randomisation restreinte, la randomisation adaptative sur cova-

riables, la randomisation adaptative sur réponses et la randomisation adaptée au covariables et aux réponses.

1.4.3.1 La randomisation complète

La procédure de randomisation la plus simple est la randomisation complète définie par :

$$\Phi_n = E(T_n | \mathcal{F}_{n-1}) = E(T_n).$$

Pour $K=2$, $T_i = (0,1)$ ou $(1,0)$ pour un tirage au sort, alors T_{11}, \dots, T_{n1} sont des variables indépendantes et identiquement distribuées selon une loi Bernoulli avec la probabilité d'être assigné au traitement 1 est donnée par :

$$\Phi_{i1} = E(T_{i1}) = \frac{1}{2}, \quad i = 1, \dots, n$$

Cette procédure est utilisée rarement dans la pratique, car la probabilité de l'instabilité de traitement dans des petits échantillons est non négligeable.

1.4.3.2 La randomisation restreinte

La procédure de la randomisation restreinte est la meilleure méthode de randomisation pour plusieurs essais cliniques, car il est souvent souhaitable d'avoir un nombre égal de patients alloués à chaque traitement. Ceci est généralement accompli en modifiant la probabilité de randomisation à un traitement en fonction du nombre de patients qui a déjà été affecté à ce traitement. Cette procédure est définie par :

$$\Phi_n = E(T_n | \mathcal{F}_{n-1}) = E(T_n | \mathcal{J}_{n-1}).$$

[Rosenberger and Lachin, 2016] ont décrit 4 procédures différentes de la randomisation restreinte. La première est la règle de répartition aléatoire, où $K=2$ et n un nombre pair, la probabilité dans ce cas que le $i^{\text{ème}}$ sujet soit assigné au traitement 1 est donnée par :

$$\Phi_{i1} = E(T_{i1} | \mathcal{F}_{i-1}) = \frac{\frac{n}{2} - N_1(i-1)}{n - (i-1)}, \quad i = 1, \dots, n.$$

La deuxième procédure est la conception binomiale tronquée qui, pour n pair et $K = 2$, est donnée par :

$$\begin{aligned}\Phi_{i1} &= \frac{1}{2}, \text{ si } \max \{N_1(i-1), N_2(i-1)\} < \frac{n}{2} \\ &= 0, \text{ si } N_1(i-1) = \frac{n}{2} \\ &= 1, \text{ si } N_2(i-1) = \frac{n}{2}\end{aligned}$$

La troisième procédure est le modèle présenté par [Efron, 1971]. À partir du vecteur $N(i)$, soit, $D_i = N_1(i) - N_2(i)$ représente le déséquilibre entre les traitements 1 et 2. Nous définissons une constante $\pi \in [0.5; 1]$, alors la procédure est définie par :

$$\begin{aligned}\Phi_{i1} &= \frac{1}{2}, \text{ si } D_{i-1} = 0 \\ &= \pi, \text{ si } D_{i-1} < 0 \\ &= 1 - \pi, \text{ si } D_{i-1} > 0\end{aligned}$$

Enfin, la dernière procédure de l'urne, design de Wei [Wei, 1978] propose une modification sur le modèle présenté par [Efron, 1971] tel que le degré de déséquilibre influence la procédure de randomisation. Une technique est d'établir un modèle d'urne dans lequel les balles dans l'urne sont étiquetées $1, \dots, K$ et chacune représente le traitement à affecter. Soit Y_n la composition de l'urne, où Y_{nj} est le nombre de balles étiquetées j après que n patients ont été assignés, $j = 1, \dots, K$. On commence par une composition d'urne initiale $Y_0 = 1$. Chaque randomisation est accomplie comme suit : une balle est retirée et remplacée, son étiquette notée, le traitement approprié est assigné et une balle de chacune des autres étiquettes $K-1$ est ajoutée à l'urne.

Alors la procédure de randomisation restreinte est présentée par :

$$\Phi_i = \frac{Y_{i-1}}{\|Y_{i-1}\|} = \frac{\Phi_{i1} - N(i-1)}{i(K-1) + 1}$$

Une procédure généralisée a été proposée par [Wei et al., 1986]. Soit $\pi(\cdot) = (\pi_1(\cdot), \dots, \pi_K(\cdot))$ une fonction continue telle qu'elle satisfait la relation suivante :

$$\frac{N_j(i-1)}{i-1} \geq \frac{1}{K}$$

Alors,

$$\pi_j\left(\frac{N(i-1)}{i-1}\right) \leq \frac{1}{K}, j = 1, \dots, K$$

Ainsi, la procédure de randomisation restreinte est donnée par :

$$\Phi_i = \pi\left(\frac{N(i-1)}{i-1}\right)$$

1.4.3.3 La randomisation adaptative sur les réponses

La randomisation adaptative à la réponse est une technique de randomisation dans laquelle l'allocation des patients aux groupes de traitement est basée sur la réponse (résultat) des patients précédents. L'objectif est de fournir aux patients un meilleur / meilleur traitement basé sur la connaissance de l'effet du traitement à ce moment-là. En conséquence, la randomisation de réponse adaptative prend la préoccupation éthique en considération. Dans ce cas les procédures de la randomisation adaptative sur réponses s'écrivent de la manière suivante :

$$\Phi_n = E(T_n | \mathcal{F}_{n-1}) = E(T_n | \mathcal{T}_{n-1}, \mathcal{X}_{n-1}).$$

1.4.3.4 La randomisation adaptative sur covariables

La randomisation adaptative sur covariables est généralement considérée pour réduire le déséquilibre des covariables entre les groupes de traitements. Ainsi, la randomisation adaptative sur covariables est également connue sous le nom de stratification adaptative. Dans ce cas les procédures de randomisation adaptative sur covariables s'écrivent de la manière suivante :

$$\Phi_n = E(T_n | \mathcal{F}_{n-1}) = E(T_n | \mathcal{T}_{n-1}, \mathcal{Z}_n).$$

La procédure de [Pocock and Simon, 1975] est une des procédures la plus utilisée dans ce cas. Soit, Z_1, \dots, Z_n le vecteur de covariables des patients $1, \dots, n$. En outre, nous supposons qu'il existe des covariables S et divisées en différents niveaux n_s , avec $s = 1, \dots, S$.

Nous définissons N_{sik} le nombre de patients dans le $i^{\text{ème}}$ niveau pour la $s^{\text{ème}}$ covariable sur le traitement k , avec $s = 1, \dots, S$, $i = 1, \dots, n_s$ et $k=1,2$.

Soit le patient $n+1$ avec un vecteur de covariable $Z_{n+1} = (r_1, \dots, r_S)$. On définit la différence entre le nombre de patients pour les traitements 1 et 2 au niveau r_s de la covariable s , $D_s = N_{sr_s1}(n) - N_{sr_s2}(n)$.

Soit $\omega_1, \dots, \omega_S$ l'ensemble des poids tel que l'agrégation pondérée $D(n) = \sum_{s=1}^S \omega_s D_s(n)$.

Le procédure d'allocation du traitement 1 avec une probabilité $\pi \in \left[\frac{1}{2}, 1\right]$ est :

$$\begin{aligned}\Phi_{i1} &= E(T_{i1}|\mathcal{T}_{i=1}, \mathcal{Z}_i) \\ &= \frac{1}{2}, \quad \text{si } D_{i-1} = 0 \\ &= \pi, \quad \text{si } D_{i-1} < 0 \\ &= 1 - \pi, \quad \text{si } D_{i-1} > 0\end{aligned}$$

1.4.3.5 La randomisation adaptée aux covariables et aux réponses

Dans la randomisation adaptée aux covariables et aux réponses, les réponses et les vecteurs de covariables du patient précédent avec le vecteur de covariable du patient actuel sont utilisés pour obtenir la probabilité de l'ajustement des effets du traitement induits par les covariables ajustée sur l'affectation au traitement. Cette procédure peut être présentée comme suit :

$$\Phi_n = E(T_n|\mathcal{F}_{n-1}) = E(T_n|\mathcal{T}_{n-1}, \mathcal{X}_{n-1}, \mathcal{Z}_n).$$

On considère le modèle de régression logistique ajusté aux covariables, défini par :

$$\text{logitPr}(X_{ij} = 1|T_{ij} = t_{ij}, Z_i = z_i) = \alpha + \beta t_{ij} + z_i \gamma' + t_{ij} z_i \delta'$$

Avec, α est l'ordonnée à l'origine, β est l'effet principal du traitement, γ est un vecteur de taille S des effets principaux des covariables, et δ est un vecteur de taille S des interactions du traitement par rapport aux covariables.

Le modèle adapté après n patients donne des estimateurs de maximum de vraisemblance $\hat{\alpha}$, $\hat{\beta}$, $\hat{\gamma}$, $\hat{\delta}$. L'odd-ratio de la covariable ajustée est basé sur les données de n patients et le vecteur de covariable du n+1 patients est donné par :

$$\frac{e^{\hat{\alpha}_n + \hat{\beta}_n + z_{n+1} \hat{\gamma}' + z_{n+1} \hat{\delta}'}}{e^{\hat{\alpha}_n + z_{n+1} \hat{\gamma}'}} = e^{\hat{\beta}_n + z_{n+1} \hat{\delta}'}$$

Nous définissons alors la procédure de randomisation sous la forme ci-dessous :

$$\Phi_{ij} = E \left(T_{ij} | \mathcal{T}_{i-1}, \mathcal{X}_{i-1}, \mathcal{Z}_{i-1} = \frac{e^{\hat{\beta}_{i-1} + z_i \hat{\delta}'_{i-1}}}{1 + e^{\hat{\beta}_{i-1} + z_i \hat{\delta}'_{i-1}}} \right)$$

1.5 Conclusion sur les méthodes adaptatives

L'utilisation de modèles adaptatifs dans les essais cliniques peut être une initiative majeure pour augmenter l'efficacité de l'essai. Comprendre différents types de designs adaptatifs et les défis associés à chaque conception en termes de logistique et des statistiques revêt une importance capitale dans la mise en place des essais.

Avant le début des essais, une étude minutieuse des avantages et des inconvénients est nécessaire pour évaluer leur utilité. La mise en place d'essais adaptatifs nécessite un support informatique en termes de gestion des essais, d'approvisionnements cliniques, d'analyses statistiques et de logiciels de simulation. Le temps supplémentaire et l'effort requis pour planifier un processus adaptatif peuvent faire gagner du temps et de l'argent à long terme.

1.6 Présentation de stratégies statistiques participant à la résolution des difficultés statistiques dans le cas des maladies rares

1.6.1 Les difficultés statistiques dans le cas des maladies rares

L'utilisation d'une valeur de p (P-value) de 0,05 pour distinguer le résultat des analyses statistiques en «significatif» et «non significatif» est presque universelle dans la recherche clinique, bien qu'il n'y ait pas de justification statistique formelle pour l'utilisation d'une valeur de 0,05 et aucune considération pour savoir si d'autres valeurs seraient plus appropriées dans des contextes particuliers de recherche clinique [Berry, 2006]. Dans l'étude des maladies rares, où les patients sont par définition difficiles à recruter, il est souvent impossible d'arriver à la taille de l'échantillon nécessaire pour atteindre la puissance habituelle de 80 %.

La randomisation dans un essai contrôlé randomisé- ERC permettra d'équilibrer les bras dans la répartition des patients selon leurs caractéristiques. Mais la randomisation ne produira pas nécessairement l'équilibre entre un ERC de taille appropriée à la question posée et les ERC de plus petites tailles. Un grand déséquilibre dans les covariables peut compromettre la validité d'un essai. Avec une petite taille d'échantillon, il existe des limites à l'application des méthodes habituelles pour la résolution du problème de déséquilibre - randomisation stratifiée sur les covariables importantes et l'ajustement par régression dans la comparaison des résultats [Brookes et al., 2001].

De nombreuses techniques d'analyse standards dépendent dans une certaine mesure sur les hypothèses d'une distribution normale (des observations, des résidus ou des estimations). Avec de petites tailles d'échantillons, ces hypothèses peuvent être difficiles à vérifier et différentes analyses, qui peuvent sembler appropriées, peuvent conduire à des résultats différents. Par exemple, dans une petite étude avec des résultats continus, un t-test et le test de Mann-Whitney peuvent aboutir à des conclusions différentes [Abrahamyan et al., 2016]. Nous présentons dans la suite trois stratégies statistiques participant à la résolution des difficultés statistiques dans le cas des maladies rares.

1.6.2 Utilisation des résultats de substitution dans les études des maladies rares

Les critères d'évaluation principale d'un essai clinique doivent être une mesure bien définie et fiable, facilement mesurable et sensible à l'intervention et cliniquement pertinente pour le patient [Aronson, 2005]. Cependant, l'identification de ces critères et la mesure de leurs changements peuvent être difficiles dans de nombreuses maladies rares, compte tenu du manque d'évidence sur la progression de la maladie, des temps de réponse thérapeutiques variables et inconnus, de la présence d'un dommage pathologique irréversible au moment du diagnostic et des maladies qui sont incurables avec un objectif thérapeutique limité au ralentissement de la progression de la maladie et/ou à la gestion des symptômes. Pour ces raisons, et généralement pour les avantages de la diminution des exigences de taille de l'échantillon, les enquêteurs peuvent préférer l'utilisation de critères de substitution - qui sont plus faciles à mesurer dans les essais de maladies rares.

Cependant, si le substitut doit être utilisé comme critère d'évaluation primaire, il est important d'établir une relation statistiquement robuste entre un résultat de substitution et le résultat clinique d'intérêt fixé dans la littérature. Plus largement encore, le critère de substitution devrait être en mesure de saisir l'effet net de l'intervention par rapport à la mesure de l'efficacité clinique [Fleming and Powers, 2012].

Les critères de substitution, tels que les biomarqueurs, qui permettent de prédire si les patients subiront des résultats cliniques d'intérêt, peuvent également être utiles dans les études des maladies rares. Mais la validation des biomarqueurs, en tant que bons substituts du résultat clinique d'intérêt, peut être difficile. Ils peuvent améliorer davantage le pouvoir statistique, si un nombre potentiellement faible de patients présente le critère d'intérêt, alors que presque tous les patients ont des valeurs mesurées du biomarqueur.

1.6.3 Utilisation des scores de Propension

Certains auteurs ont suggéré l'utilisation de certaines méthodes avancées, comme les scores de propension, pour lutter contre la confusion dans les études de santé des maladies rares [Cole et al., 2011], [Sun and Garrison, 2012]. Lorsqu'on compare les patients traités pour une maladie rare particulière à des patients atteints de la même maladie, mais qui ne sont pas traités, des confusions se produiront si les déterminants de la réception d'un traitement par rapport à un autre sont également des facteurs de risque. Souvent, de nombreux facteurs de confusion peuvent être présents. Les scores de propension réduisent la dimensionnalité de la confusion dans les études d'observation en résumant tous les fac-

teurs de confusion potentiels en un seul score scalaire [Rosenbaum and Rubin, 1983]. Cet outil est particulièrement utile dans les études où il y a peu d'événements par rapport au nombre de facteurs de confusion, ce qui est une caractéristique des maladies rares [Cepeda et al., 2003].

Bien que les scores de propension puissent faciliter l'ajustement pour de nombreux facteurs de confusion potentiels, en revanche ni les scores de propension ni la modélisation traditionnelle de la régression des résultats ne peuvent pallier la confusion due à des variables non mesurées.

1.6.4 Planifier une méta-analyse prospective

Même si les chercheurs utilisent une de ces conceptions innovantes ou des méthodes adaptatives présentées dans ce chapitre, les essais individuels des patients peuvent ne pas être capables d'obtenir une puissance statistique suffisante pour rejeter l'hypothèse nulle à l'aide d'un seuil conventionnel ($\alpha = 0,05$). Une solution est d'augmenter α , comme cela a été fait dans l'essai alternatif de conception de l'Itraconazole par [Gallin et al., 2003]. Une autre solution consiste à mener l'étude et d'incorporer les résultats dans une méta-analyse planifiée préalablement [Gagne et al., 2014].

Pour la recherche sur des maladies à faible prévalence ou incidence, comme les maladies rares, le nombre d'individus atteints à tout moment peut rendre impossible de mener même un ERC multicentrique qui pourrait distinguer de manière fiable les interventions [Edwards et al., 1997] [Knapp, 1996].

Des fausses conclusions négatives peuvent être tirées et l'analyse de puissance sera impossible d'élucider l'erreur [Goodman, 1994]. Si les chercheurs prévoient explicitement que les résultats d'un petit essai soient disponibles pour être inclus dans une méta-analyse prospective, les risques d'erreur de types *I* et *II* de fausses conclusions peuvent être évités [Simes, 1995]. Les méta-analyses prospectivement conçues sont moins sensibles aux problèmes liés aux méta-analyses traditionnelles, car les méthodes d'étude des composants de la méta-analyse peuvent être synchronisées à l'avance. On part de K études dont la méta-analyse est non "productive". On construit une étude supplémentaire en tenant compte des paramètres de K études afin d'espérer que la nouvelle méta-analyse soit conclusive.

Par conséquent, seules les méta-analyses conçues prospectivement peuvent justifier les risques pour les participants dans les essais de faible puissance parce qu'ils fournissent une assurance suffisante que le résultat d'une étude contribuera éventuellement à des connaissances importantes. Bien qu'un essai multicentrique peut contribuer à des connaissances importantes de manière similaire à une méta-analyse prospective, mais pour cela il faut

que les cliniciens aient accès à un nombre suffisant de patients pendant la conduite de l'essai, ce qui n'est pas le cas dans les études des maladies rares.

La méta-analyse prospective est une méthode qui a été utilisée au cours des dernières années par des expérimentateurs dans les maladies cardiovasculaires et la leucémie chez les enfants. Il peut s'agir d'une méthodologie utile lorsque, par exemple, de grandes tailles d'échantillons sont nécessaires pour assurer une puissance statistique adéquate, mais les essais à grande échelle ne sont pas réalisables. Une alternative est que les chercheurs effectuent leur propre étude localement et collaborent avec les chercheurs d'études similaires, en organisant les résultats à combiner à la fin de chaque essai (permettant ainsi une méta-analyse continue, cumulative et prospective). Cela permet aux chercheurs individuels de maintenir une certaine autonomie et, en même temps, de planifier de manière appropriée la méta-analyse. Une autre situation où cela peut être bénéfique est lorsque deux ou plusieurs essais portant sur la même question clinique commencent. Une fois que les essais similaires sont identifiés, les chercheurs peuvent collaborer et planifier prospectivement la combinaison leurs résultats dans une méta-analyse.

Nous présentons ci-dessous les éléments d'un protocole de planification d'une méta-analyse prospective [Halpern SD et al., 2002].

Objectif :

- Définir les hypothèses/objectifs spécifiques.

Critères pour examiner les études pour cette méta-analyse

- Critères d'éligibilité pour la conception de l'essai (par exemple, les exigences de randomisation, la durée minimale de suivi).
- Critères d'éligibilité pour la population de patients.
- Critères d'éligibilité pour chaque intervention et comparateur.
- Information sur la spécification de ses critères principaux et secondaires.
- Détails des sous-groupes.

Recherche des méthodes pour identifier les études :

- Décrire les efforts faits pour identifier les essais en cours.

Collecte et analyse de données :

- Liste des détails des essais identifiés pour l'inclusion dans la méta-analyse prospective.
- Les essais ne devraient être inclus que si leurs résultats étaient inconnus au moment où ils ont été identifiés et ajoutés à la méta-analyse prospective
- Détails de la taille de l'échantillon, du calcul de la puissance de cette méta-analyse, des analyses intermédiaires et des analyses de sous-groupes.
- Gestion des données (données à collecter, format requis, etc).
- Analyse statistique

1.6.5 Méthodes bayésiennes

Les méthodes bayésiennes sont recommandées dans la littérature pour résoudre le problème de manque de puissance dans les essais cliniques des maladies rares [Lilford et al., 1995], [Tan et al., 2003a], [Goodman and Sladky, 2005], [Greenhouse and Seltman, 2005], car l'approche bayésienne est idéale pour s'adapter à l'information accumulée au cours d'un essai, ce qui permet d'obtenir de plus petits essais informatifs et de recevoir un meilleur traitement.

L'approche bayésienne utilise toutes les données disponibles - de l'essai et d'autres sources - pour calculer les probabilités qu'un traitement particulier soit efficace. Ces probabilités peuvent alors être appliquées à la pratique clinique.

Les méthodes bayésiennes sont utilisées pour la conception et l'interprétation des essais cliniques menés avec des tailles d'échantillons limités. Les informations préalables sur les paramètres d'un modèle bayésien sont résumées par les distributions *a priori* déterminées soit à partir de l'opinion d'experts soit à partir d'une combinaison de l'opinion d'experts

associés à des données connexes. Ces données pourraient être sous la forme de données générées dans une population de référence, ou plus généralement des données sur un paramètre de mesure ou sur un médicament dont le mécanisme d'action est similaire au nouveau médicament.

Nous présenterons les principes de cette méthode et son application dans le cas des maladies rares dans le Chapitre 2.

1.7 Discussion

Cette revue de la littérature des méthodes a été axée sur des articles qui ont mentionné «maladie rare» dans un champ consultable. En raison du grand nombre de maladies rares, nous n'étions pas en mesure de rechercher des applications de méthodes novatrices liées à chaque maladie spécifique. En outre, notre examen visait à fournir un aperçu général des méthodes non traditionnelles qui ont été proposées ou appliquées à l'étude des maladies rares.

S'il y a d'autres méthodes non traditionnelles qui existent et qui pourraient être applicables à des maladies rares, mais n'ont pas encore été discutées dans une publication dans les bases de données que nous avons cherchées, nous ne les avons pas identifiées. En outre, notre revue est destinée à mieux faire connaître la disponibilité et l'utilisation de méthodes novatrices pour étudier des essais épidémiologiques et cliniques dans les maladies rares, cette recherche est aussi destinée à fournir un examen technique de ces méthodes. Nous avons retiré 21 méthodes qui peuvent être utilisées dans les essais cliniques pour étudier les interventions thérapeutiques dans les maladies rares.

Les investigateurs qui étudient les maladies rares ont essayé de faire face au petit groupe de participants potentiels aux essais. Certains ont proposé des ajustements aux essais randomisés traditionnels. En se concentrant sur les patients à haut risque, une idée était de réduire la taille de l'échantillon et la durée de l'étude.

Certains chercheurs ont tenté de réduire la taille de l'échantillon en abordant multiples options de traitements dans le cadre d'une étude factorielle. Une autre manière de réduire les exigences de taille d'échantillon dans les études de maladies rares concerne la sélection du critère d'évaluation en utilisant les résultats des variables continues ou d'un marqueur de substitution. Les critères de substitution, comme les biomarqueurs, qui permettent de prédire si les patients présenteront les résultats cliniques d'intérêt peuvent également être utiles. Ils peuvent améliorer la puissance statistique lorsque le nombre de patients dans une étude est potentiellement réduit.

De plus, de nombreuses variantes de randomisation adaptatives et de modèles séquentiels sont applicables à l'étude des maladies rares, car ils peuvent réduire la taille de l'échantillon nécessaire dans le contexte d'essais conventionnels. En outre, certains modèles adaptatifs peuvent également augmenter la probabilité de recevoir le traitement le plus efficace, ce qui peut encourager l'adhésion à l'essai. Même si les enquêteurs utilisent une adaptation innovante de l'un de ces modèles d'essais traditionnels dans l'étude d'un

traitement des maladies rares, des essais individuels de la santé des patients peuvent ne pas être capables d'atteindre une puissance suffisante pour rejeter l'hypothèse nulle en utilisant un seuil fréquentiste classique ($\alpha = 0.05$). Une solution est d'augmenter α , une autre solution est de procéder une étude de faible puissance et d'intégrer les résultats dans une méta-analyse prospective planifiée. Une troisième option est d'incorporer les résultats dans un cadre de travail bayésien.

Certains de ces modèles peuvent également réduire les besoins de recrutement par rapport aux essais randomisés contrôlés des groupes parallèles classiques, par exemple, les essais croisés et les essais de type N-of-1. Mais ces essais ne peuvent généralement pas être utilisés pour étudier les traitements qui ont des effets curatifs ou des conditions qui évoluent rapidement. Beaucoup de maladies rares sont des maladies chroniques qui progressent lentement et inexorablement au cours du temps.

En conclusion, sur la base des articles étudiés, il y a peu de recherche structurée sur les possibilités statistiques visant à améliorer les preuves de petites études d'échantillons. En dehors de cela, nous avons observé un manque d'étude systématique de la conception et des méthodes statistiques pour les études sur les maladies rares, de sorte que d'autres recherches seraient souhaitables. Cependant, les recommandations générales sont difficiles dans le domaine des maladies rares, car les situations et les tailles d'échantillons varient énormément. L'utilisation d'une approche bayésienne peut augmenter la robustesse de l'information à partir de petits essais et peut être utilisé pour aider à concevoir et fournir une justification pour des tels essais. Cette approche utilise toutes les données disponibles, de l'essai et d'autres sources pour calculer les probabilités qu'un traitement particulier soit efficace. Ces probabilités peuvent ensuite être appliquées à la pratique clinique pour arriver à une conclusion ou bien à une meilleure compréhension de la problématique.

Nous proposons à la fin du travail de nouvelles méthodes bayésiennes et leurs applications à la maladie de Horton.

1.8 Avantages et Inconvénients des méthodes

Le tableau des avantages et limites de chaque méthodes est présenté dans l'annexe.

Chapitre 2

Introduction aux méthodes bayésiennes et leurs applications aux maladies rares

2.1 Introduction

2.1.1 Introduction du chapitre

Ce chapitre est organisé en deux parties. La première partie introduit la méthodologie de l'approche bayésienne (les paragraphes 2.2 à 2.8). Ensuite nous présentons l'application de cette approche aux maladies rares (les paragraphes 2.9 à 2.11).

2.1.2 Comment se définit la statistique bayésienne

La statistique bayésienne est une approche permettant de déduire la probabilité d'un événement à partir des probabilités d'autres événements déjà évalués. Dans les essais cliniques, les méthodes statistiques fréquentistes peuvent utiliser des informations qui proviennent d'études précédentes au stade de la conception [Buckley, 2008]. À l'étape de l'analyse des données, les informations de ces études sont considérées comme un complément, mais ne font pas partie de l'analyse formelle. En revanche, l'approche des méthodes statistiques bayésiennes utilise le théorème de Bayes pour combiner l'information *a priori*

avec des informations en lien avec l'étude considérée pour en déduire une loi *a posteriori*. L'approche bayésienne fournit un paradigme complet à la fois pour l'inférence statistique et la prise de décision en situation d'incertitude. Elle peut être le support d'un système axiomatique, et donc fournir une méthodologie générale cohérente. Les méthodes bayésiennes peuvent se rapprocher de la décision fréquentiste dans des cas très particuliers et "extrêmes". Elles résolvent certaines des difficultés rencontrées par les méthodes statistiques classiques, et étendent l'applicabilité des méthodes statistiques en choisissant une loi *a priori* et une fonction de coût. En particulier, les méthodes bayésiennes donnent la possibilité d'incorporer des hypothèses scientifiques dans l'analyse (par le biais de la distribution *a priori*) et peuvent être appliquées à des problèmes dont la structure est trop complexe pour que les méthodes classiques puissent s'appliquer.

L'utilisation de techniques bayésiennes offre plusieurs avantages importants par rapport à l'approche statistique classique. Premièrement, contrairement aux statistiques classiques, l'analyse bayésienne permet de calculer directement la probabilité qu'un traitement soit bénéfique conditionnellement aux mesures observées. Deuxièmement, la statistique bayésienne permet au chercheur d'incorporer n'importe quelle information *a priori* à l'analyse des résultats expérimentaux. Le paradigme bayésien est basé sur une interprétation de la probabilité comme étant une mesure de l'incertitude à la fois conditionnelle et rationnelle.

L'inférence statistique sur une quantité d'intérêt est décrite comme la modification de l'incertitude sur sa valeur- intervalle de crédibilité- à la lumière des mesures observées, et le théorème de Bayes spécifie précisément comment cette modification doit être faite. La situation particulière, souvent rencontrée dans les rapports scientifiques, est la prise de décision quand la seule information acceptable est celle qui peut être déduite de données disponibles. Dans ce cas, grâce à des lois *a priori*, les méthodes bayésiennes permettent de prendre une décision [Robert, 2007], [Berger, 1985b]. L'approche bayésienne peut être présentée comme une généralisation de l'approche classique : les paramètres ne sont plus des valeurs fixes inconnues, mais des variables aléatoires dont il faut spécifier la distribution. La distribution donnée au moment de la modélisation est dite *a priori*. L'intégration de l'information apportée par les données se fait par le calcul de la distribution dite *a posteriori* et qui n'est autre que la distribution conditionnelle - calculée par la formule du théorème de Bayes - aux valeurs prises par les données. Les hypothèses scientifiques sont généralement exprimées par des distributions de probabilité pour les données scientifiques observables. Ces distributions de probabilités dépendent de quantités inconnues appelées paramètres. L'utilisation des méthodes bayésiennes se justifie pour plusieurs raisons, et leurs applications apparaissent dans divers domaines [Robert, 2007], [Berger, 1985a]. Di-

vers théorèmes fondamentaux montrent que pour prendre des décisions cohérentes dans un contexte d'incertitude, alors l'approche bayésienne est un support approprié. D'autres soulignent des problèmes logiques avec des méthodes fréquentistes qui ne se posent pas dans le cadre bayésien [Samaniego, 2010], [Apostolakis, 1981].

D'autre part, les probabilités *a priori* sont intrinsèquement subjectives - votre information préalable est différente de la mienne - et de nombreux statisticiens voient cela comme un inconvénient fondamental de la statistique bayésienne. Les défenseurs de l'approche bayésienne soutiennent que c'est inévitable, et que les méthodes fréquentistes entraînent aussi des choix subjectifs. Cela a été la source de discorde entre les partisans fondamentalistes des deux paradigmes statistiques durant les 50 dernières années. En revanche, ce sont plus les avantages pragmatiques de l'approche bayésienne qui ont alimenté sa forte croissance au cours des 20 dernières années, et qui sont la raison de son adoption dans une variété de domaines en croissance rapide [Samaniego, 2010], [Apostolakis, 1981].

2.2 Comment planifier une analyse bayésienne dans les essais cliniques.

Étape 1 :

Les principes de base d'un bon design d'essai sont les mêmes dans les cas bayésiens et fréquentistes d'essais. Les parties d'un plan expérimental complet comprennent :

- les objectifs de l'essai,
- le critère principal à évaluer,
- les conditions sous lesquelles l'essai va être mené,
- la population étudiée et l'analyse statistique prévue.

Étape 2 : La sélection d'un critère de jugement principal :

Le critère principal est en général une mesure d'efficacité utilisée pour explorer une hypothèse. Le critère principal doit répondre à certaines caractéristiques :

- cliniquement pertinent,
- directement observable,
- important pour le patient.

Par exemple, un critère principal peut être une mesure : mortalité, morbidité, ou qualité de vie.

Étape 3 : Collecter d'autres informations importantes : les covariables

Les covariables qui sont les facteurs de confusion, sont caractéristiques des patients de l'étude et peuvent influencer sur les résultats. Il y a beaucoup de techniques statistiques (bayésiennes et fréquentistes) pour ajuster les variables. L'ajustement de covariables est particulièrement important dans toute situation où un certain degré d'équilibre de la covariable n'est pas assuré par la randomisation. Par exemple dans un essai bayésien dans lequel d'autres essais sont utilisés comme information préalable, si les ajustements ne sont pas faits entre les variables des différents essais, l'analyse peut être biaisée.

L'ajustement des covariables est souvent utilisé pour réduire la variabilité, ce qui conduit à une analyse plus puissante.

Étape 4 : Choisir une comparaison : Contrôle

Pour faciliter l'évaluation des résultats d'un essai clinique, il est recommandé d'utiliser un comparateur, ou groupe contrôle, comme référence.

Les types des groupes contrôles qui peuvent être considérés sont :

- contrôles concomitants : les deux groupes "contrôle et témoin" sont suivis de manière similaire
- self-control : chaque patient est son propre contrôle,
- contrôles historiques : un groupe de patients recevant un nouveau traitement peut être comparé au groupe des patients qui, par exemple, durant l'année précédant l'introduction du nouveau traitement, ont été traités avec l'ancien traitement.

Une autre façon de caractériser le type de contrôle est de distinguer entre les contrôles qui sont traités avec une thérapie efficace (contrôles actifs) vs contrôles qui ne reçoivent aucun traitement ou sont traités avec un dispositif fictif (le groupe placebo, contrôle inactif).

Les méthodes bayésiennes sont particulièrement utiles dans les essais contrôlés visant à démontrer qu'un nouveau dispositif est non seulement non-inférieur au contrôle actif, mais est également supérieur au contrôle inactif ou fictif.

Les méthodes bayésiennes peuvent être utiles pour combiner les contrôles historiques avec les contrôles concomitants. Les contrôles historiques servent d'informations *a priori* pour les contrôles concomitants.

2.3 L'approche bayésienne

2.3.1 L'approche bayésienne

Supposons A et B deux événements aléatoires, tels que $\mathbb{P}(A)$ et $\mathbb{P}(B) \neq 0$. La probabilité de B conditionnellement à la réalisation de A, et la probabilité de A conditionnellement à la réalisation de B sont exprimées par :

$$\mathbb{P}(B|A) = \frac{\mathbb{P}(A, B)}{\mathbb{P}(A)} \text{ et } \mathbb{P}(A|B) = \frac{\mathbb{P}(B, A)}{\mathbb{P}(B)}$$

Avec $\mathbb{P}(A, B) = \mathbb{P}(B, A)$ = probabilité que les deux événements A et B aient lieu simultanément.

Remplaçons $\mathbb{P}(B, A)$ par $\mathbb{P}(A, B) = \mathbb{P}(A|B)\mathbb{P}(B)$, on peut en déduire la **formule de Bayes** qui est la relation entre les deux probabilités conditionnelles $\mathbb{P}(A, B)$ et $\mathbb{P}(B, A)$

présentées par l'équation suivante :

$$\mathbb{P}(B|A) = \frac{\mathbb{P}(A|B)\mathbb{P}(B)}{\mathbb{P}(A)} \quad (2.1)$$

Le modèle bayésien peut se représenter par le graphique suivant :

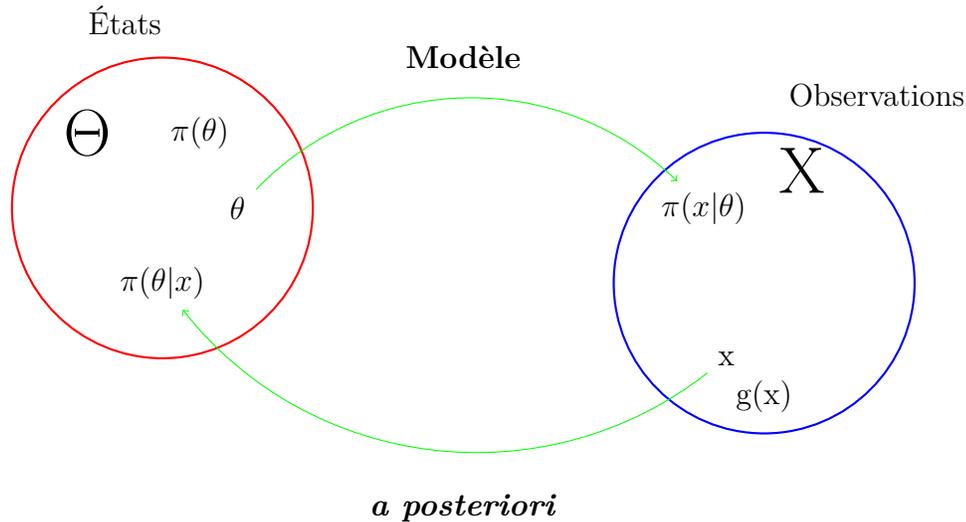


Figure 2.1 – Principe de l'analyse bayésienne

Les différents termes du graphique ci-dessus sont définis de la manière suivante.

Soit Θ l'espace des paramètres $\theta = (\theta_1, \theta_2, \dots, \theta_n)$ et \mathbb{X} l'ensemble des observations tel que l'on dispose d'un échantillon aléatoire de n observations notées $x = (x_1, x_2, \dots, x_n)$ issues d'un modèle statistique des distributions possibles Π_θ .

Vraisemblance : Considérons la densité de cet échantillon $f(x|\theta)$ comme une fonction de θ , cette fonction appelée de vraisemblance de θ en x est notée $L(\theta; x) = \pi(x|\theta)$. $\pi(x|\theta)$ est la probabilité des n observations $x = (x_1, x_2, \dots, x_n)$ conditionnellement à la valeur des paramètres θ .

Loi a priori :

La connaissance *a priori* sur le paramètre θ est entachée d'incertitude et s'exprime, à travers une loi de probabilité nommée loi *a priori* notée π de $\pi(\theta)$. La notation *a priori*

exprime le fait qu'elle a été établie avant d'avoir des informations sur les valeurs d'observations x et donc cette loi peut permettre de traduire les connaissances que l'on a avant l'expérience.

Elle peut être issue de l'opinion des experts ou bien établie sur la base de l'analyse d'autres données similaires ou "voisines" ou bien sur l'avis personnel du statisticien [Albert et al., 2012], [O'Hagan et al., 2006] [Ayyub, 2000].

On distingue deux types de lois *a priori* : les lois informatives et les lois non informatives que l'on expliquera brièvement dans le paragraphe suivant.

Loi *a posteriori* : L'information sur θ est mise à jour grâce au théorème de Bayes qui permet de prendre en compte l'information apportée par les observations. Ainsi on obtient la loi *a posteriori* de θ . Lorsque les nouvelles données " x " deviennent disponibles, les informations qu'elles contiennent concernant les paramètres du modèle sont exprimées dans la "probabilité", qui est proportionnelle à la distribution des données observées compte tenu des paramètres du modèle, exprimées par la vraisemblance $\pi(x|\theta)$.

Cette information est ensuite combinée avec l'information *a priori* pour produire une distribution de probabilité actualisée appelée "distribution *a posteriori*", sur laquelle est basée toute inférence bayésienne. Le théorème de Bayes, une identité élémentaire de la théorie des probabilités, indique comment la mise à jour de la connaissance sur la base des informations se fait mathématiquement (figures 2.1 et 2.2) :

$$\pi(\theta|x) = \frac{\pi(\theta) \times \pi(x|\theta)}{\int_{\Theta} \pi(\theta) \times \pi(x|\theta)d\theta} = \frac{\pi(\theta)\pi(x|\theta)}{g(x)} \quad (2.2)$$

La quantité $g(x) = \int_{\Theta} \pi(\theta) \times \pi(x|\theta)d\theta$ est la loi marginale de x également appelée constante d'intégration.

Comme $g(x)$ est indépendante de θ donc on peut écrire la loi *a posteriori* de la manière suivante :

$$\pi(\theta|x) \propto \pi(\theta) \times \pi(x|\theta) \quad (2.3)$$

En théorie, la distribution *a posteriori* est toujours disponible, mais dans les modèles complexes réalistes, les calculs analytiques requis sont souvent impossibles.

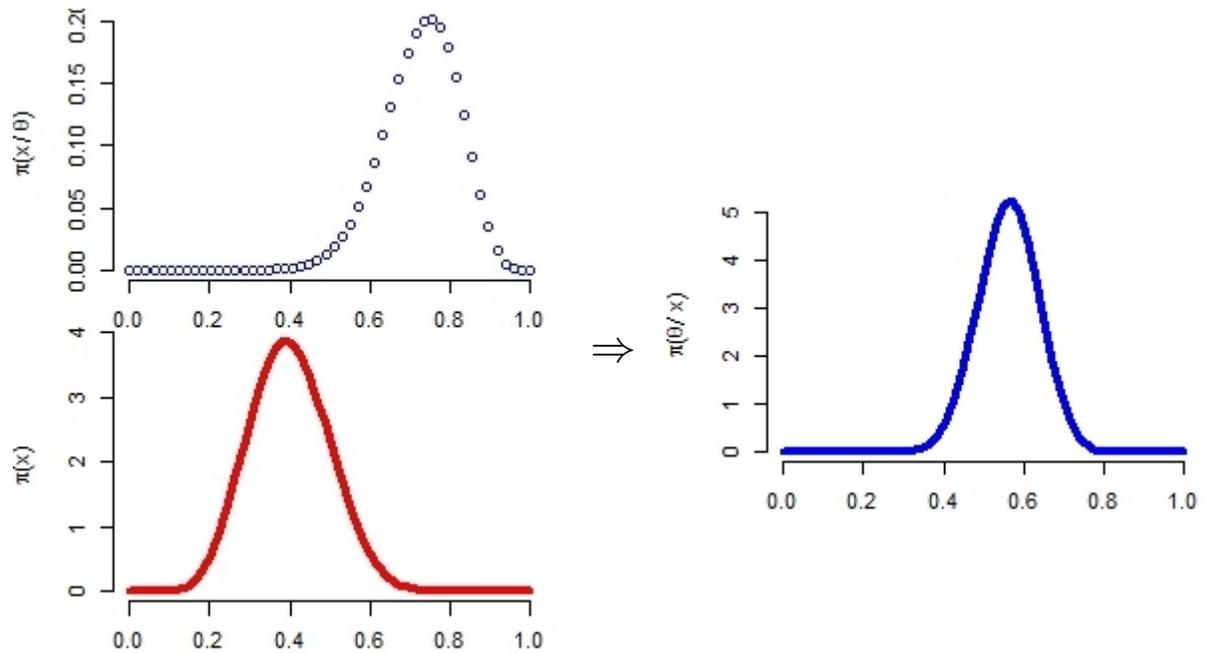


Figure 2.2 – Vraisemblance \propto a priori \propto a posteriori

Exemple 2.3.1. On présente un exemple sur l'application de la méthode bayésienne. Soit x le nombre de succès observés et θ la probabilité de ce succès. Donc le modèle convenable à cette situation est la loi binomiale. Ce qui veut dire que la probabilité d'observer x succès conditionnellement à θ s'écrit :

$$\pi(x|\theta) = C_n^\theta \theta^x (1 - \theta)^{n-x} \quad (2.4)$$

En statistique bayésienne les lois a priori conjuguées du paramètre θ de la loi binomiale sont présentées par les lois de la famille bêta de paramètres α et β . Donc, l'expression de la loi a priori de θ est :

$$\pi(\theta) = \frac{1}{B(\alpha, \beta)} \theta^{\alpha-1} (1 - \theta)^{\beta-1} \quad (2.5)$$

D'après la formule de Bayes (équation 2.3), la loi a posteriori $\pi(\theta|x)$ est proportionnelle au produit entre la vraisemblance et la loi a priori, donc :

$$\pi(\theta|x) \propto \theta^{x+\alpha-1}(1-\theta)^{n-x+\beta-1} \quad (2.6)$$

Alors, la loi a posteriori est encore une loi bêta,

$$\pi(\theta|x) \hookrightarrow B(x+\alpha, n-x+\beta).$$

On suppose maintenant l'exemple d'une pièce normale, on souhaite connaître la probabilité θ de tomber sur pile. On procède 20 lancers, on obtient l'échantillon $\{P, P, F, F, P, P, P, P, P, F, F, F, F, P, P, F, P, F, P, P\}$, soit 12 pile et 8 face. La probabilité d'obtenir ce résultat s'exprime en fonction de θ par :

$$\pi(x|\theta) = \pi(x=12) = C_{20}^{12} \theta^{12} (1-\theta)^8 \quad (\text{équation 2.4}).$$

On vérifie que la vraisemblance est maximale pour $\theta = 0.6$ (figure 2.3).

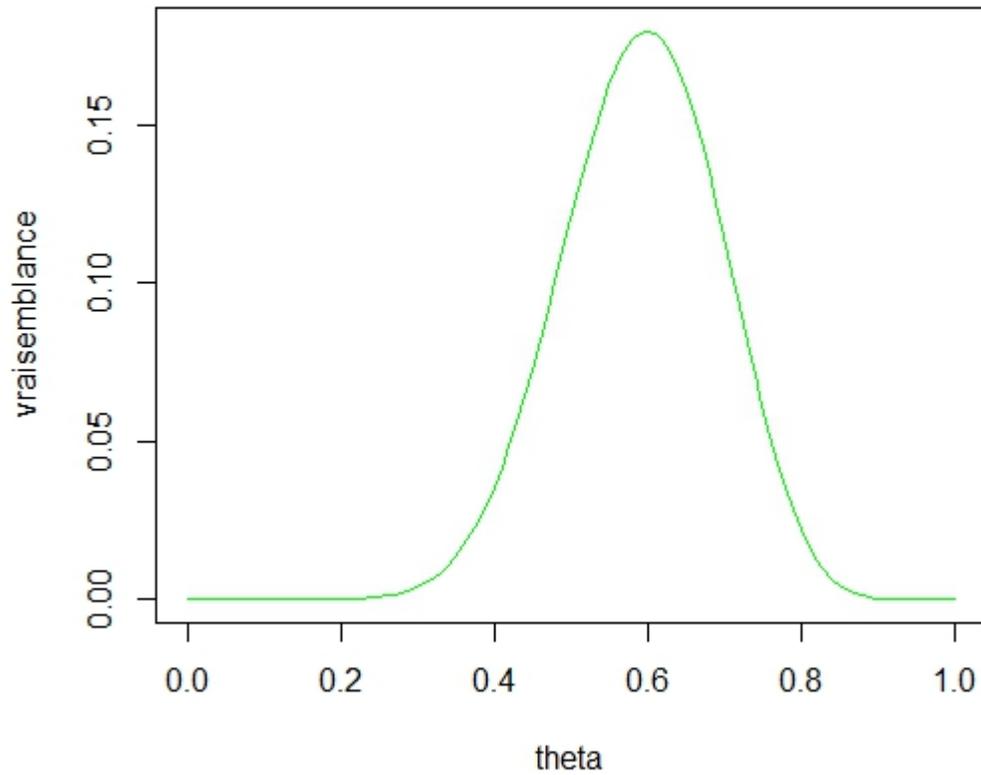


Figure 2.3 – Vraisemblance

On veut connaître la probabilité θ de tomber sur pile pour une pièce donnée. Nous supposons la croyance a priori que toutes les valeurs de θ sont équiprobables (a priori non informative). Donc nous prenons comme loi a priori une loi bêta, telle que :

$$\pi(\theta) \propto B(1, 1)$$

En faisant les calculs de l'équation 2.6, nous obtenons la distribution a posteriori :

$$\pi(\theta|x) \propto \theta^{12}(1 - \theta)^8$$

Donc la loi a posteriori est une loi bêta de paramètres : 13 et 9.

Nous représentons graphiquement la vraisemblance et les densités a priori et a posteriori (figures 2.4) :

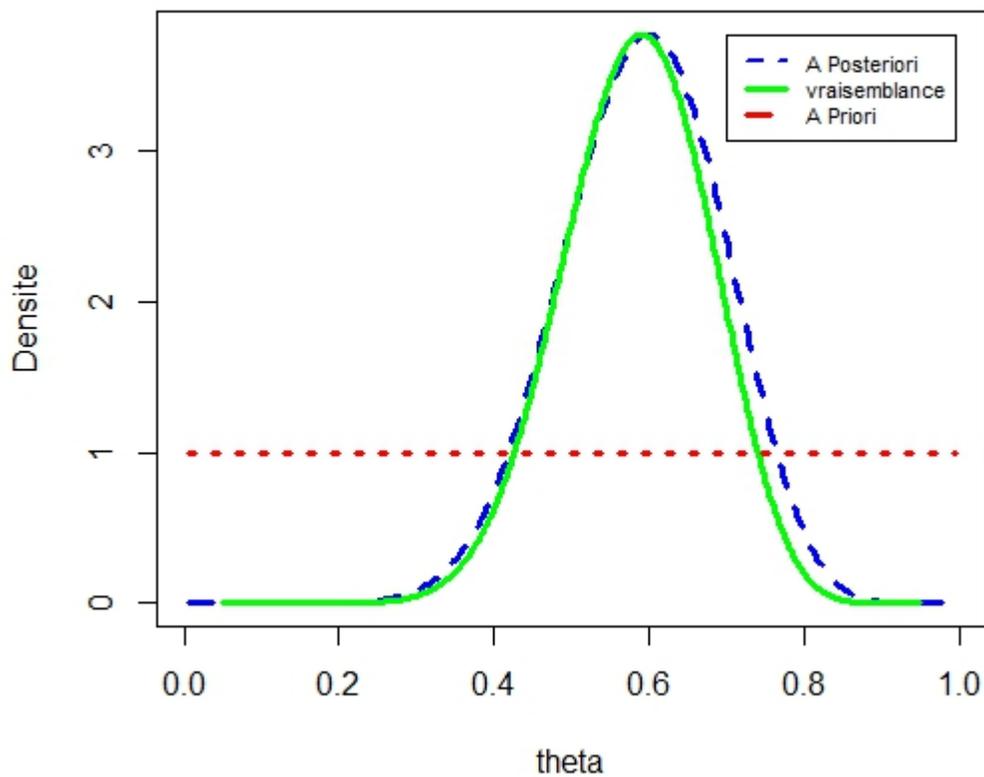


Figure 2.4 – Représentation graphique de la vraisemblance, les lois a priori et a posteriori de l'exemple

On remarque que dans ce cas comme l'a priori est non informatif, la distribution a posteriori du paramètre sera proche de la vraisemblance.

2.4 Loi *a priori* informative et non informative

Une distribution *a priori* informative est une loi non dominée par la vraisemblance et qui a un impact sur la distribution *a posteriori*. Si une distribution *a priori* domine sur la vraisemblance, il s'agit clairement d'une loi *a priori* informative. Ces types de distributions doivent être spécifiées avec précaution dans la pratique.

Une distribution *a priori* informative privilégie pour certaines valeurs de la quantité d'intérêt. Ces préférences sont généralement basées sur des études antérieures. Si l'information *a priori* pour une étude est basée sur beaucoup plus de patients que ceux inscrits à l'étude, la distribution *a priori* peut être trop informative.

Le jugement de ce qui constitue une distribution trop informative est une décision au cas par cas. Dans ce cas, certaines modifications de l'étude et du plan d'étude peuvent être justifiées.

Lorsque les informations *a priori* sur le modèle sont trop vagues ou peu fiables, le choix de la distribution *a priori* est très difficile. D'autres raisons peuvent expliquer l'absence d'une distribution *a priori* bien définie. En outre, les exigences d'objectivité peuvent obliger le statisticien à fournir une réponse avec le moins possible d'entrées subjectives, afin de baser l'inférence sur le modèle d'échantillonnage seul. Cependant, en gardant à l'esprit les fondements bayésiens des critères d'optimalité fréquentistes, il semble préférable de suivre une approche bayésienne, en utilisant un *a priori* dérivé du modèle en tant qu'outil technique. Le manque de préférence entre les valeurs ou le manque d'information peuvent être représentés par une distribution *a priori* non informative. Quand aucune information préalable est disponible, on a recours aux distributions *a priori* non informatives qui sont fréquemment utilisées dans les essais adaptatifs bayésiens [Robert, 2007], [Berger, 1985b]. En plus de la nécessité d'une entrée minimale des données subjectivement, les distributions *a priori* conjuguées peuvent en effet être considérées comme un point de départ pour établir les lois *a priori* basées sur une entrée préalable limitée dont l'imprécision peut être évaluée par des distributions *a priori* supplémentaires. L'utilisation des lois conjuguées facilite le calcul de la loi *a posteriori*.

Définition 2.4.1. Une famille \mathbb{F} de distributions de probabilité sur Θ est considérée comme conjuguée pour une fonction de vraisemblance $f(x|\theta)$ si, pour tout $\pi \in \mathbb{F}$, la distribution *a posteriori* $\pi(\theta|x) \in \mathbb{F}$

En effet, l'intérêt principal de l'utilisation d'une famille conjuguée devient plus évident lorsque \mathbb{F} est aussi petit que possible et paramétré. D'ailleurs, lorsque \mathbb{F} est paramétré, le passage de la distribution *a priori* à l'*a posteriori* est réduit à une mise à jour des

paramètres correspondants.

Cette propriété seule peut expliquer pourquoi les *a priori* conjugués sont si populaires, car les distributions *a posteriori* sont toujours calculables. Donc, le praticien induit directement la forme de son estimateur dès qu'il a choisi sa loi *a priori*.

Donc dans le tableau suivant (Table 2.1) nous avons présenté les lois *a priori* conjuguées

| $\mathbf{f}(\mathbf{x}/\theta)$ | $\pi(\theta)$ | $\pi(\theta x)$ | $\mathbb{E}(\theta x)$ |
|--|---|---|---|
| $\mathcal{N}(\theta, \sigma^2)$ | $\mathcal{N}(\mu, \tau^2)$ | $\mathcal{N}(x/\sigma^2 + \mu/\tau^2, \frac{1}{1/\sigma^2 + 1/\tau^2})$ | $x/\sigma^2 + \mu/\tau^2$ |
| $\mathcal{G}(n, \theta)$ | $\mathcal{G}(\alpha, \beta)$ | $\mathcal{G}(\alpha + n, \beta + x)$ | $(\alpha + n)/(\beta + x)$ |
| $\mathcal{B}(n, \theta)$ | <i>Bêta</i> (α, β) | <i>Bêta</i> $(\alpha + n, \beta + x)$ | $(\alpha + n)/(\alpha + n + \beta + x)$ |
| $\mathcal{P}(\theta)$ | $\mathcal{G}(\alpha, \beta)$ | $\mathcal{G}(\alpha + n, \beta + 1)$ | $(\alpha + x)/(\beta + 1)$ |
| $\mathcal{M}_k(\theta_1, \dots, \theta_k)$ | $\mathcal{D}(\alpha_1, \dots, \alpha_{1k})$ | $\mathcal{D}(\alpha_1 + x_1, \dots, \alpha_k + x_k)$ | |

Table 2.1 – Exemple de lois conjuguées

qui sont utiles comme approximations de loi *a priori* réelles. Cependant, lorsqu'aucune information préalable n'est disponible, il est impossible de justifier le choix des lois *a priori* sur une base subjective et les hyperparamètres de la loi *a priori* conjuguée ne peuvent être déterminés. Ce qui est nécessaire dans de telles situations est un *a priori* non-informatif, c'est-à-dire un *a priori* qui ne contient aucune information sur θ . Par exemple, en faisant le test entre deux hypothèses simples, l'*a priori* qui donne la probabilité $\frac{1}{2}$ à chacune des hypothèses est clairement non informatif.

Deux approches sont couramment utilisées pour le choix de cet *a priori* non informatif [Berger, 1985b] :

- le principe de raison insuffisante (ou bien principe de Laplace) est basé sur l'équiprobabilité des événements élémentaires et semble donc être assez fiable,
- l'approche de Jeffrey qui permet d'adopter une approche ne nécessitant pas de structure d'invariance d'*a priori* invariante. Cette approche est basée sur l'information de Fisher $I(\theta)$, telle que la distribution *a priori* dans le cas unidimensionnel est donnée par :

$$\pi^*(\theta) \propto I^{1/2}(\theta)$$

$$\text{avec } I(\theta) = \mathbb{E}_\theta \left[\left(\frac{\partial \log f(x|\theta)}{\partial(\theta)} \right)^2 \right]$$

D'autre part, lorsque θ est un paramètre multidimensionnel, l'*a priori* non informatif de Jeffrey est alors défini par :

$$\pi^*(\theta) \propto [\det(I(\theta))]^{1/2}$$

tel que la matrice d'information de Fisher est définie par $I_{i,j}(\theta) = -\mathbb{E}_\theta \left[\frac{\partial^2 \log f(x|\theta)}{\partial(\theta_i) \partial(\theta_j)} \right]$

Nous présentons dans la section suivante un type important de loi *a priori* qui concerne les *a priori* basés sur le modèle hiérarchique et qui permet une spécification *a priori* permettant par exemple de mélanger des *a priori* informatifs et non informatifs.

2.5 Modèles hiérarchiques

2.5.1 Introduction

L'approche bayésienne peut être considérée comme utile pour la modélisation à multiples facettes et les phénomènes non linéaires tels que ceux rencontrés dans les maladies rares.

Le manque de données au niveau individuel, couplé à la volonté de prendre en compte des différences individuelles expose à des difficultés dans l'analyse des données des maladies rares.

Les modèles hiérarchiques bayésiens sont la combinaison de deux éléments :

1. un modèle écrit sous forme hiérarchique,
2. des estimations qui utilisent les méthodes bayésiennes.

Un modèle hiérarchique est celui qui est écrit de façon modulaire, ou en termes de sous-modèles.

Il est utilisé lorsque les données sont structurées en groupe, exemple : spatialement, temporellement, démographiquement, etc., et que différents paramètres (reliés) sont utilisés pour chaque groupe. Dans les modèles hiérarchiques, on peut parler d'interchangeabilité à chaque niveau de groupe. Par exemple, on suppose que deux traitements sont appliqués, dans des essais randomisés séparés, à des patients de différents pays. Par la suite, s'il n'y avait pas d'autres informations disponibles, il devient raisonnable de traiter chaque patient de chaque pays comme échangeable et aussi de traiter les résultats des différents pays comme échangeables.

En pratique, il serait logique d'inclure par exemple, comme variables explicatives au niveau de la ville, toutes les informations pertinentes que nous avons sur chaque ville, ainsi

que les variables explicatives au niveau individuel, puis les distributions conditionnelles montrant que ces variables explicatives seront échangeables.

2.5.2 Théorie

Un modèle hiérarchique bayésien est un modèle statistique bayésien, $(f(x|\theta), \pi(\theta))$, tel que la distribution *a priori* $\pi(\theta)$ est décomposée en distributions conditionnelles

$$\pi_1(\theta|\theta_1), \pi_2(\theta_1|\theta_2), \dots, \pi_n(\theta_{n-1}|\theta_n)$$

et une distribution marginale $\pi_{n+1}(\theta_n)$ telle que

$$\pi(\theta) = \int_{\Theta_1 \times \dots \times \Theta_n} \pi_1(\theta|\theta_1) \pi_2(\theta_1|\theta_2) \dots \pi_{n+1}(\theta_n) d\theta_1 \dots d\theta_{n+1}$$

Les paramètres θ_i sont appelés des hyper paramètres de niveau i ($1 \leq i \leq n + 1$).

Un aspect particulièrement attrayant des modèles hiérarchiques est qu'ils permettent le conditionnement à tous les niveaux, et cette décomposition simple présente la distribution *a posteriori* de θ ([Lemme 2.5.1](#))

Lemme 2.5.1. *Dans le cas où on a :*

$$\theta|\theta_1 \propto \pi_1(\theta|\theta_1), \quad \theta_1 \propto \pi_2(\theta_1)$$

La loi *a posteriori* de θ est

$$\pi(\theta|x) = \int_{\Theta_1} \pi(\theta|\theta_1, x) \pi(\theta_1|x) d\theta_1$$

avec,

$$\begin{aligned} \pi_1(\theta|x, \theta_1) &= \frac{f(x|\theta) \pi_1(\theta|\theta_1)}{m_1(x|\theta_1)} \\ \text{avec } m_1(x|\theta_1) &= \int_{\Theta} f(x|\theta) \pi_1(\theta|\theta_1) d\theta, \\ \text{et } \pi(\theta_1|x) &= \frac{m_1(x|\theta_1) \pi_2(\theta_1)}{m(x)} \\ \text{avec } m(x) &= \int_{\Theta_1} m_1(x|\theta_1) \pi_2(\theta_1) d\theta_1 \end{aligned}$$

2.6 La théorie de décision en analyse bayésienne

Soit D l'ensemble de décisions i.e l'ensemble de toutes les décisions possibles. Nous suivons ici les notations de [Robert, 2007] et [Berger, 1985a] avec X l'espace des observations et Θ l'espace des paramètres.

Notons $\delta : X \rightarrow D$ un estimateur de θ , qui à chaque observation x , associe une décision $\delta(x)$. La plupart du temps, l'espace des décisions pour l'estimation ponctuelle correspond à l'espace des paramètres, ou à une transformation de cet espace, $h(\Theta)$. L'inférence statistique consiste à prendre une décision d permettant d'estimer le paramètre θ ou une fonction du paramètre θ , $h(\theta)$.

Définition 2.6.1. *Une fonction de coût est une fonction définie de la manière suivante :*

$$L : \Theta \times D \rightarrow [0, +\infty)$$

La fonction de coût $L(\theta, d)$ évalue l'erreur associée à la décision d quand le paramètre prend la valeur θ .

Dans le cadre de la théorie de décision, on est en mesure d'évaluer cette décision.

La théorie de la décision bayésienne est fondée sur le coût moyen a posteriori.

Ainsi, au lieu d'utiliser le critère de comparaison fréquentiste (équation 2.8) qui est une fonction de θ qui ne permet pas de comparer tous les décisions et estimateurs :

$$R(\theta, \delta) = E_{\theta} [L(\theta, \delta(x))] \quad (2.7)$$

$$= \int_X L(\theta, \delta(x)) f(x|\theta) dx \quad (2.8)$$

On utilise le coût moyen a posteriori, qui est la base dans la théorie de décision de l'approche bayésienne, tel que l'on intègre la fonction de coût dans l'espace Θ puisque θ est inconnu, au lieu d'intégrer sur l'espace X étant donné que x est connu. Elle repose sur le coût a posteriori.

$$\begin{aligned} \rho(\pi, d|x) &= E^{\pi} [L(\theta, d)|x] \\ &= \int_{\Theta} L(\theta, d) f(\theta|x) d\theta \\ &= \int_{\Theta} L(\theta, d) \pi_x(d\theta) \quad \text{tel que } \pi_x(d\theta) = f(\theta|x) \\ &= E_{\pi_x} [L(\theta, d)] \end{aligned}$$

On remarque que le coût moyen a posteriori est fonction de x , contrairement au critère

de comparaison fréquentiste qui dépend de θ . On définit le risque bayésien $r(\pi, \delta)$ comme le risque fréquentiste moyenné sur les valeurs de θ en fonction de leur distribution a posteriori.

On obtient pour chaque estimateur un nombre réel fourni par $r(\pi, \delta)$ (Théorème 2.6.1), et non une fonction de θ comme dans l'approche fréquentiste. Ce risque mesure le coût moyen associé à l'estimateur δ . Donc ceci revient à minimiser $\rho(\pi, d|x)$.

La minimisation de ce risque est l'un des buts de la théorie de la décision statistique.

Théorème 2.6.1. *Un estimateur minimisant le risque bayésien $r(\pi, \delta)$ peut être obtenu en choisissant pour $x \in X$, la valeur $\delta(x)$ qui minimise le coût a posteriori, $\rho(\pi, \delta|x)$ puisque :*

$$r(\pi, \delta) = \int_X \rho(\pi, \delta|x)m(x)dx$$

Preuve: La preuve est présentée dans l'Annexe du chapitre 2 □

Le théorème est intéressant en pratique puisque si l'on connaît la loi de X, faire la moyenne sur l'ensemble des valeurs de x est une perte d'informations, il suffit de moyenner le coût a posteriori $\rho(\pi, \delta|x)$.

On remarquera qu'en général l'estimateur bayésien est unique et que si la loi est impropre (voire Définition 2.6.2), on peut construire un estimateur de Bayes généralisé qu'on peut noter δ^π qui minimise $r(\pi, \delta)$. Pour chaque $x \in X$, on a l'estimateur donné par :

$$r(\pi, \delta^\pi) = \inf_{\delta \in D} r(\pi, \delta) < \infty \tag{2.9}$$

Donc $r(\pi, \delta^\pi)$ est appelé le risque bayésien. Pour obtenir la valeur de $\inf_{\delta \in D} r(\pi, \delta)$, il suffira de minimiser une grandeur qui ne dépend plus que des données. Ce calcul permet donc d'arriver à des estimations admissibles.

On dira que δ_1 est meilleur que δ_2 au sens du risque bayésien si :

$$r(\pi, \delta_1) \leq r(\pi, \delta_2)$$

Il est donc toujours possible de comparer deux estimateurs de δ au sens du risque de Bayes, et on dit que δ est admissible s'il n'existe pas de meilleure règle de décision.

Exemple 2.6.1. ([Berger, 1985a]) On suppose $D = \Theta \subset \mathbb{R}$, et $L(\theta, \delta) = |\theta - \delta|$, donc

$$\begin{aligned} \rho(\pi, \delta|x) &= \int_{\Theta} |\theta - \delta| d\pi(\theta|X) \\ &= \int_{-\infty}^{\delta} (\delta - \theta)\pi(\theta|X)d(\theta) + \int_{\delta}^{\infty} (\delta - \theta)\pi(\theta|X)d(\theta) \end{aligned} \quad (2.10)$$

Comme $\delta \rightarrow \rho(\pi, \delta|x)$ est convexe et dérivable presque partout, il suffit de déterminer les points critiques :

$$\frac{\partial \rho(\pi, \delta|x)}{\partial \delta} = \int_{-\infty}^{\delta} \pi(\theta|X)d(\theta) - \int_{\delta}^{\infty} \pi(\theta|X)d(\theta) = 0,$$

donc, $\int_{-\infty}^{\delta} \pi(\theta|X)d(\theta) = \int_{\delta}^{\infty} \pi(\theta|X)d(\theta)$,

$$\iff P^{\pi}(\theta \leq \delta|X) = P^{\pi}(\theta \geq \delta|X).$$

Alors, l'estimateur bayésien $\delta^{\pi}(X)$ est la médiane de $\pi(\theta|X)$

Beaucoup d'autres notions sont importantes en théorie de la décision. Nous allons définir les plus importantes. La notion de minimaxité permet de minimiser le coût moyen dans le cas le moins favorable.

Définition 2.6.2. soit la loi a priori impropre de θ notée par :

$$\int \pi(\theta)d\theta = +\infty.$$

L'intérêt d'introduire une telle notion est de permettre d'introduire l'estimateur bayésien généralisé.

On propose une loi a priori impropre $\pi(\theta)$, telle que l'intégrale

$$\int_{\theta} L(\theta, x)\pi(\theta)d\theta \text{ est convergente.}$$

Donc on appelle estimateur de Bayes généralisé de θ , la moyenne de la loi a posteriori

$$\pi(\theta|x) = \frac{L(\theta, x)\pi(\theta)}{\int_{\theta} L(\theta, x)\pi(\theta)d\theta}$$

Définition 2.6.3. Le principe du minimax consiste à se protéger contre le plus mauvais cas, donc on peut calculer $\sup_{\theta} R(\theta, \delta)$.

On dira qu'une règle de décision est minimax si elle minimise $\sup_{\theta} R(\theta, \delta)$.

Soit :

$$\left\{ \begin{array}{l} \delta * \text{ telle que :} \\ \inf_{\delta \in D} \sup_{\theta} R(\theta, \delta) = \sup_{\theta} R(\theta, \delta *) \end{array} \right.$$

En résumé, on choisit la règle de décision qui rend le risque maximum le plus petit. S'il n'y a pas de données, on appelle cette règle de décision une "action minimax".

2.7 Choix du modèle

Une caractéristique clé de sélection de modèles bayésiens lorsque les modèles ont des dimensions différentes et des paramètres non communs est que les résultats sont généralement très sensibles au choix des lois *a priori* pour les paramètres non communs et, contrairement au scénario pour l'estimation, cette sensibilité ne disparaîtra pas lorsque la taille de l'échantillon devient plus grande.

La loi *a priori* a moins de poids quand l'échantillon est grand que quand il est petit.

En raison de la difficulté d'évaluer des lois *a priori* subjectives pour de nombreux modèles, il a fallu développer des lois *a priori* "classiques" ou "objectives" pour la sélection du modèle. Le mot objectif signifie tout simplement que ce ne sont pas des lois *a priori* subjectives, et donc qu'elles sont choisies en se basant sur les modèles envisagés.

Notations : soit y un vecteur de taille n d'un des modèles :

$$M_0 : f_0(y|\alpha), \dots, M_i : f_i(y|\alpha, \beta_i), i = 1, 2, \dots, N - 1$$

tel que α et β_i sont les paramètres inconnus du modèle. M_0 va être nommé le modèle nul et est inclus dans tous les modèles considérés.

Dans le cadre du modèle nul(H_0), la loi *a priori* est $\pi_0(\alpha)$, dans le cadre du modèle M_i et sans perte de généralité, on exprime la loi *a priori* comme suit :

$$\pi_i(\alpha, \beta_i) = \pi_i(\alpha)\pi_i(\beta_i|\alpha)$$

Notons que le paramètre α se trouve dans tous les modèles, et habituellement α est désigné comme paramètre commun, les β_i sont désignés comme les paramètres spécifiques du modèle.

En supposant que l'un des modèles envisagés est vrai, la loi *a posteriori* de chaque modèle M_i peut être écrite sous la forme :

$$P(M_i|y) = \frac{B_{i0}}{1 + (\sum_{j=1}^{N-1} B_{j0}P_{j0})}$$

tel que pour $j \neq 0$, $P_{j0} = \frac{P(M_j)}{P(M_0)}$, avec $P(M_j)$ est la probabilité *a priori* du modèle M_j , et B_{j0} est le facteur bayésien du modèle M_j à M_0 défini par :

$$B_{j0} = \frac{m_j(y)}{m_0(y)}, \text{ avec } m_j(y) = \int f_j(y|\alpha, \beta_i)\pi_j(\alpha, \beta_j)d\alpha d\beta_j$$

et $m_0(y) = \int f_0(y|\alpha)\pi_0(\alpha)d\alpha$ est la loi marginale du modèle M_j et M_0 correspondant aux densités *a priori* du modèle $\pi_j(\alpha, \beta_j)$ et $\pi_0(\alpha)$.

2.8 Critères pour une sélection objective de la loi *a priori*

Les lois *a priori* des paramètres β_i doivent être propres, car elles apparaissent dans le numérateur du facteur bayésien B_{i0} , donc si on utilise une loi *a priori* impropre la constante arbitraire correspondante ne ne serait pas annulée, ce qui implique que B_{i0} serait mal défini.

Similairement, les lois *a priori* vagues (qui sont propres) ne peuvent pas être utilisées pour déterminer le B_{i0} .

2.9 Méthodes bayésiennes dans le cas de maladies rares.

Les méthodes bayésiennes sont recommandées par plusieurs auteurs et dans le cas de ces méthodes il faut introduire une loi informative *a priori* sur l'effet inconnu du traitement : ceci se fait à partir soit d'expertises ou de données rétrospectives [Lilford et al., 1995], [Goodman and Sladky, 2005], [Tan et al., 2003b], [Hampson et al., 2014].

Les données sur les témoins issues de RCT historiques peuvent être synthétisées à partir de méta-analyses bayésiennes et utilisés pour en déduire une loi *a priori* des paramètres de nouveaux essais en tenant compte de l'incertitude et de l'hétérogénéité entre traitements [Neuenschwander et al., 2010], [Schmidli et al., 2013], [Gsteiger et al., 2013].

Ces auteurs ont proposé deux grands types d'analyse (qui s'appliquent aux études sur les enfants), qui peuvent être associés et qui dépendent des connaissances acquises soit à partir d'experts soit à partir des données issues des études chez les adultes d'où :

- Les lois *a priori* sont déduites des méta-analyses de données disponibles des adultes, à condition que la physiopathologie et/ou le traitement soient voisins.
- Les lois *a priori* sont déduites d'informations issues chez les adultes associées à celles publiées chez les enfants en supposant que les effets des traitements entre les deux groupes sont échangeables.

Nous présentons dans la suite les travaux les plus pertinents sur ce sujet.

2.10 Perturbation de la loi *a priori* par les informations provenant d'autres études notamment chez les adultes.

Introduction

[Goodman and Sladky, 2005] dans leur article démontrent comment les méthodes bayésiennes peuvent être utilisées pour incorporer une information *a priori* de l'efficacité du traitement chez des adultes pour faire un essai randomisé de non-infériorité comparant des immunoglobulines intraveineuses (IgIV) vs la plasmaphérèse chez les enfants.

Un modèle basé sur l'*a priori* a été construit et est équivalent à un essai sur 72 enfants montrant l'équivalence entre les thérapies. Un modèle bayésien est construit sur le hasard ratio supposé suivre une loi normale. Il conduit à une taille de l'échantillon de 160 enfants,

avec une analyse pré postérieure démontrant un type d'erreur de première espèce de 5 % et une puissance de 77 %.

Cette étude illustre la construction d'un modèle basé sur l'*a priori* qui permettrait avec l'information des adultes d'augmenter formellement les données auprès des enfants afin de minimiser une expérimentation pédiatrique inutile. La discussion pour le choix des lois *a priori* adaptées à de tels designs est à la fois un élément nécessaire et souhaitable pour les essais bayésiens.

Exemple : le syndrome de Guillain-Barré

Le syndrome de Guillain-Barré évolue en 4 phases :

- événement de départ,
- période de progression,
- période en plateau,
- période de récupération.

L'histoire du syndrome Guillain-Barré est connue chez les adultes. Chez les adultes et les enfants, la pathologie est identique excepté pour le critère d'évaluation et la récupération (récupération plus rapide chez les enfants et plus complète chez les adultes). Les deux traitements initiaux sont : la plasmaphérèse et l'immunoglobulines intraveineuses.

Évidence de l'efficacité des traitements chez les adultes.

- Au milieu des années 1980 deux RCT comparant la plasmaphérèse vs placebo ont donné le même hasards ratio sur la variable temps de récupération [[Plasmapheresis, 1985](#)], [[Efficiency, 1987](#)],
- le temps médian de la reprise de la marche dans le bras placebo de ces deux essais différait considérablement : 111 jours dans la première étude et 85 dans la deuxième,
- deux larges RCT ont ensuite été menés chez les adultes comparant plasmaphéreses vs IgIV, un en 1992 [[Van der Meché and Schmitz, 1992](#)] et le deuxième en 1997 [[Plasma, 1997](#)],
- le premier a montré de faibles preuves pour un petit bénéfice des IgIV par rapport à la plasmaphérèse, et le second présente des courbes de récupération à la marche superposables .

Évidence de l'efficacité des traitements chez les enfants.

- La qualité des données chez les enfants est inégale.

- 3 cas d'essais non contrôlés [Paradiso et al., 1999], [Korinthenberg and Mönting, 1996], [Graf et al., 1999].
- 4 cas d'essais d'enfants traités par la plasmaphérèse comparés avec des témoins historiques [Jansen et al., 1993], [Lamont et al., 1991], [Khatiri et al., 1990], [Epstein and Sladky, 1990].
- temps de récupération : autour des grandes valeurs
- les résultats sont cohérents globalement avec ceux des adultes.
- le HR de la méta-analyse est de 0.58 (IC=[0.32 – 1]) en faveur de la plasmaphérèse. Il est à peu près identique à ceux des adultes, mais basé sur des données de qualité inégale.

Protocole de l'étude :

La question posée est l'efficacité de la plasmaphérèse vs IgIV chez les enfants. Les deux traitements sont considérés comme standard et on ne peut pas éthiquement comparer ces traitements au placebo.

Une question se pose : comment un essai randomisé qui inclut les résultats des adultes pourrait-il être une base pour le processus de décision pour les IgIV (vs plasmaphérèse) ? On choisit un essai de non-infériorité justifié par l'équivalence d'efficacité des 2 traitements observés chez les adultes (i.e les IgIV peuvent être préférées à la plasmaphérèse, car elles sont beaucoup moins pourvoyeuses de complications et moins "coûteuses" que la plasmaphérèse, même si elles sont un peu moins efficaces).

On estime le temps médian de reprise de la marche pour la plasmaphérèse est de 24 jours. On choisit une marge de non-infériorité de 7 jours, ce qui implique que tant que la durée moyenne du retard à la marche est inférieure à 31 jours le traitement est supposé acceptable, avec un hazards ratio de : $\frac{31}{24} = 1.3$

Construction de la loi *a priori*

Le HR chez les adultes pour la plasmaphérèse vs IgIV est de 0.91 (IC=[0.75 – 1.1]). Dans les essais pédiatriques, il n'est généralement pas justifié d'appliquer les résultats obtenus chez les adultes. Les données des adultes ne sont pas classiquement un support pour constituer la base de la loi *a priori* chez les enfants. Toutefois, se basant sur les caractéristiques cliniques et physiopathologiques de la maladie d'autant que les données des enfants sont de mauvaise qualité, on constate que les données des adultes devraient être informatives chez l'enfant.

Pour combiner la loi *a priori*, les résultats chez l'adulte ont été modifiés de 2 façons :

1. Pour éliminer un biais en faveur des IgIV, on choisit pour la loi *a priori* un HR de 1 (à la place de 0.91).
2. Pour refléter l'incertitude de l'application des données de l'adulte à l'enfant, la quantité d'information fournie par les données des adultes est réduite par un facteur 2.

Les propriétés de cette loi *a priori* sont alors étudiées pour voir si elles sont cohérentes avec l'opinion des experts.

L'ajustement mathématique produit une courbe de loi *a priori* centrée sur la valeur 1 avec un intervalle de crédibilité de [0.6 – 1.67] en supposant une loi normale pour le hasards ratio.

Ceci conduit à 2.5% de probabilité *a priori* que les IgIV augmentent le temps de reprise de la marche à 38j (soit 2 semaines de plus que pour la plasmaphérèse) et à 14% de la loi *a priori* que ce temps soit de plus d'une semaine.

Les opinions correspondent au point de vue des experts. Un autre avis pour étudier cette loi *a priori* est de supposer une loi normale pour le log HR et une variance de $1/d$ (où d est le nombre d'événements).

Le choix de la loi *a priori* donne une probabilité de 86% de non-infériorité clinique. C'est-à-dire si le critère de décision est d'avoir 95% de certitude de non-infériorité, à partir de cette loi *a priori*, ceci conduit à avoir un facteur de Bayes

$$\frac{\frac{\Pi_1(\theta/x)}{\Pi_1(\theta)}}{\frac{\Pi_0(\theta/x)}{\Pi_0(\theta)}} = \frac{95/5}{86/14} = 3.1$$

Ceci est un seuil modéré, reflétant le fait que la loi *a priori* permet de recourir à moins de données pour cet essai particulier.

Caractéristiques opérationnelles :

Les caractéristiques opérationnelles définissent comment cet essai se comporte face à une variété d'hypothèses et sont calculées par simulation.

Le critère d'arrêt est la probabilité de non-infériorité $\geq 95\%$. Il donne les probabilités en pourcentage d'arrêter l'essai quand $P(HR \geq 1.3|données) \geq 0.95$ sous différentes hypothèses à propos de la vraie différence du temps médian de marche. Une augmentation de 7j de marche augmente le HR jusqu'à 1.3. Le plan de monitoring pour examiner les données est après chaque bloc de 40 patients jusqu'à un maximum de 160 patients.

Sous l'hypothèse que les traitements sont équivalents, ou bien que l'efficacité des IgIV est inférieure à la marge d'infériorité(7 jours).

Ainsi, la puissance pour la taille d'échantillon maximale = 67 % et l'erreur $\alpha= 12\%$.

L'analyse bayésienne consiste à calculer la probabilité d'arrêter l'étude à chaque point de monitoring (40,80,120,160) pour chaque hasards ratio possible avec pour hypothèses

$$H_0 : HR > 1.3$$

et

$$H_1 : HR \leq 1.3$$

et de faire la moyenne de ces probabilités en utilisant les poids des distributions de probabilité *a priori*.

Pour N=160 la probabilité d'arrêter l'essai est de 5% sous l'hypothèse H_0 et la puissance est de 77 % (arrêter à la fin de l'essai revient à dire qu'on est à 95% sûr que les IgIV sont non-inférieures).

Les nombres peuvent être utiles pour calculer la taille attendue, qui est 104 sous l'hypothèse nulle et 156 sous l'hypothèse alternative.

Conclusions

L'élément essentiel de ce protocole est de montrer comment les méthodes bayésiennes permettent de structurer le problème de telle sorte de fournir des éléments d'informations pour un échange entre les régulateurs, les intervenants et les experts. Ce design oblige à disposer des explicitations sur les questions critiques que les approches statistiques standard omettent, ou traitent de façon ad hoc.

Il faut tenir compte de plusieurs éléments :

- Le seuil du degré de certitude dans la conclusion (95 % est un choix)
- La qualité des données chez les adultes et leur pertinence chez les enfants.
- La qualité et la quantité des données chez les enfants et leur pertinence pour le protocole.
- Comment représenter la connaissance *a priori* basée sur les données de l'adulte et de l'enfant.

Pour revenir à la question initiale de savoir si nous pouvons appliquer une approche bayésienne à l'expérimentation pédiatrique lorsque les données *a priori* proviennent d'une population de patients adultes, cet exemple montre que le raisonnement bayésien par lui-

même ne peut pas répondre à cette question. L'approche bayésienne doit tenir compte du fait que les processus biologiques impliqués dans la maladie et le traitement peuvent être considérés comme connexes dans les deux populations. Notre exemple montre que, une fois que le jugement est rendu, les méthodes bayésiennes offrent un formalisme puissant qui permet à de nombreux jugements subtils réalisés par des concepteurs d'essais et des régulateurs d'être explicites, et qui apporte des informations pour une discussion raisonnée.

Il en résulte une conception de l'essai, compte tenu de la connaissance préalable de la maladie et son traitement chez les enfants et les adultes, qui limite de façon rationnelle le nombre de sujets nécessaire à l'expérimentation chez les enfants en tenant compte de la faisabilité et de l'éthique [[Buckley, 2008](#)].

2.11 Méthodes bayésiennes dans le cas des maladies rares

Des échantillons de petites tailles peuvent être notés dans les essais cliniques, soit, car la maladie est rare, soit par insuffisance de recrutement. Dans les deux cas, il faut penser différemment l'analyse statistique, car les approches plus traditionnelles peuvent être problématiques (chapitre 1). Les méthodes statistiques bayésiennes sont de plus en plus utilisées dans la recherche clinique, car l'approche bayésienne est idéale pour s'adapter à l'information accumulée au cours d'un essai, ce qui permet d'obtenir de plus petits essais informatifs et de recevoir un meilleur traitement.

Les essais cliniques visent à recueillir des preuves concernant un effet non biaisé du traitement, mais, plutôt que d'essayer de fournir une réponse définitive en utilisant les tests d'hypothèses, une vision alternative est de considérer les essais comme un moyen de réduire l'incertitude de la taille de l'effet du traitement.

Si l'on part du principe qu'il y a une incertitude importante à l'égard de cette quantité inconnue, pour les données provenant même d'un petit nombre de patients dans un essai clinique bien conçu il conviendra de prendre des mesures pour réduire cette incertitude. Cette amélioration de l'information aidera les cliniciens dans leurs prises de décision relative au traitement dont les patients ont besoin [David J. Spiegelhalter, 2003]. Cette approche statistique particulière se base sur l'utilisation de l'approche bayésienne. C'est le point de vue et la méthodologie proposée par [Lilford et al., 1995], l'un des premiers articles pour discuter de la conception des essais dans les maladies rares.

[Lilford et al., 1995] dans leur article proposent et discutent l'approche bayésienne comme une alternative appropriée aux essais avec petits échantillons mêmes si une réponse définitive ne peut pas aboutir dans la pratique. Ils précisent que les méthodes bayésiennes utilisent toutes les données disponibles pour calculer les probabilités qui peuvent être extrapolées directement dans les essais cliniques, et proposent une méthodologie pour créer une distribution *a priori* des données existantes. Ils proposent la stratégie bayésienne comme solution pour réduire le niveau d'incertitude. L'approche bayésienne ne donne pas une valeur exacte de l'effet clinique, mais la probabilité que cette valeur appartienne à un intervalle particulier. De plus, cette approche fournit des probabilités des effets du traitement qui s'appliquent directement au prochain patient semblable à ceux qui ont été traités lors d'un essai terminé ou en cours. Cet article propose une méthodologie pour calculer ces probabilités sur la base des données observées et d'une distribution *a priori*

des données existantes. Dans ce cas, la distribution *a priori* représente habituellement les attentes des cliniciens avant l'essai.

[Billingham et al., 2012] ont soutenu cette approche bayésienne proposée par [Lilford et al., 1995]. Elle permet à une information externe et subjective sur la taille de l'effet du traitement et exprimée en terme de distribution *a priori*, d'être combiné à des preuves apportées par d'autres essais pour permettre d'approcher la distribution *a posteriori* pour la taille de l'effet traitement.

Cette approche garantit que l'essai réduit l'incertitude quant à l'effet du traitement d'un niveau qui existe déjà. Par exemple, s'il existe des preuves *a priori* relativement fortes selon lesquelles un traitement est efficace, un essai de petite taille qui conforterait ce résultat pourrait être utile pour changer la pratique clinique. En revanche, la situation devient plus complexe si les données de l'essai et les preuves *a priori* se contredisent.

Un avantage clé supplémentaire de l'approche bayésienne est que les résultats d'un essai peuvent être exprimés en termes de probabilités de l'effet traitement. Par exemple, compte tenu des données probantes et des données d'un essai, il y a 70% de chances que le traitement réduise réellement le risque de décès d'au moins 10 % (c'est-à-dire un HR <0,9).

Dans les études de petite taille, ce type de rapport pourrait être utilisé pratiquement par les cliniciens et permettrait la prise de décisions de traitement fondées sur des données probantes, alors qu'un résultat non significatif du test d'hypothèse serait simplement considéré comme non concluant ou, au pire, la preuve d'aucun effet de traitement.

Suite à la proposition de Lilford et ses collègues, une stratégie a été développée par Tan et ses collègues [Tan et al., 2003a] [Tan et al., 2008] pour concevoir des essais afin d'évaluer les interventions dans les cancers rares, en particulier en termes de temps de survie comme critère de jugement. Ils suggèrent de rechercher dans la littérature toutes les preuves relatives à un essai proposé, et même d'inclure des études où il y a des tentatives similaires comme le type de la maladie, le traitement, les points extrêmes, et qui inclut tous les niveaux de preuves depuis les essais contrôlés randomisés contrôlés jusqu'aux études de cas.

Ces preuves peuvent alors être combinées avec une distribution *a priori* de l'effet du traitement avec des poids attribués par rapport à la pertinence, la validité et la précision.

En principe, cette idée est raisonnable, mais dans la pratique, une telle approche est problématique, car des stratégies de recherche générales peuvent produire un grand nombre de papiers potentiellement pertinents, mais dans les maladies rares, il est peu probable que l'un d'entre eux soit une preuve de haut niveau, car la plupart des articles sont des études de cas, donc l'examen et l'extraction des données nécessitent beaucoup de temps

et l'estimation des rapports de risque à partir de ces études sans comparaison directes de traitement n'est pas simple. Plus important encore, si cette preuve est potentiellement biaisée, alors la distribution de probabilité *a priori* ne serait pas crédible.

Compte tenu de cette difficulté à produire un *a priori* fondée sur des preuves et le fait que de nombreux cliniciens trouvent qu'il est difficile d'accepter de trouver l'*a priori* en se basant sur les croyances subjectives, nous devons considérer les solutions alternatives. En fait, l'approche bayésienne peut encore être appliquée en utilisant une distribution *a priori* non informative. Ceci est effectivement une distribution de probabilité uniforme qui reflète le fait que toutes les tailles de l'effet du traitement ont la même probabilité parce qu'il n'y a aucune preuve qui montre le contraire [Billingham et al., 2012]. Une analyse avec ce type d'*a priori* assure que la distribution **a posteriori** pour l'effet du traitement est totalement dominée par les données de l'essai. Techniquement, la distribution coïncide avec la vraisemblance, qui est une fonction de probabilité qui montre à quel point les données supportent toutes les valeurs possibles de l'effet traitement. Lorsque la vraisemblance est combinée avec une distribution *a priori* non informative, on parle souvent d'une «vraisemblance standardisée» et une telle approche pourrait être appelée une analyse bayésienne fondée sur la vraisemblance [David J. Spiegelhalter, 2003]. La raison pour laquelle une telle approche est encore utile, c'est que, comme spécifié précédemment, il permet d'exprimer les résultats en termes de probabilités directes de la taille de l'effet de traitement dans un certain intervalle, mais cette fois-ci en se fondant uniquement sur les résultats de l'essai. Ce type d'approche peut être efficace pour maximiser la valeur de l'information d'un essai qui n'a pas réussi à recruter. En ce qui concerne les maladies rares, si une telle analyse est prévue, une taille d'échantillon peut être choisie et assure que la distribution de probabilité *a posteriori* présente un niveau d'incertitude acceptable qui permettra des décisions cliniques. Cette approche utilisant une «vraisemblance standardisée» a déjà été suggérée comme une approche utile pour présenter les résultats des essais aux cliniciens [Burton, 1994], non seulement pour les petits échantillons, mais comme une analyse complémentaire des approches traditionnelles dans les essais de toute taille. Bien que le concept soit attrayant, la non-information n'est pas nécessairement simple [Hughes, 1993].

En outre, il est préférable d'utiliser une approche simple, appelée analyse conjuguée, pour estimer la distribution *a posteriori*, telle que proposée par Tan et ses collègues [Tan et al., 2003a] [Tan et al., 2008]. Ils suggèrent dans leurs articles un système de score pour mettre en commun les informations connexes des études antérieures pour voir comment cela peut se combiner avec des scénarios théoriques pour aider à modéliser la réalisation d'un RCT de petite taille.

Essentiellement, cela signifie que la distribution de probabilité postérieure est estimée simplement en combinant les valeurs des paramètres qui définissent les distributions *a priori* avec les données de l'essai (la vraisemblance), pondérées en fonction de la quantité d'information apportée par chacune. Les conclusions seront alors issues des distributions *a posteriori*. Les inférences basées sur les distributions *a posteriori* obtenues dans l'approche proposée fournissent une description complète de l'amélioration de l'état des connaissances disponibles en donnant un intervalle de crédibilité.

À long terme, la randomisation dans un essai clinique produira des groupes de patients équilibrés, mais avec des échantillons de petite taille, il existe de fortes chances de déséquilibre en termes de facteurs pronostiques potentiels et, par conséquent, ceux-ci doivent être ajustés dans l'analyse par modélisation statistique. Ainsi, une approche plus sophistiquée pour estimer les distributions *a posteriori* de l'effet du traitement peut être requise, par exemple en utilisant des méthodes de Monte-Carlo par chaînes de Markov [David J. Spiegelhalter, 2003].

Enfin, [Goodman and Sladky, 2005] dans leur article démontrent comment les méthodes bayésiennes peuvent être utilisées pour incorporer l'information *a priori* de l'efficacité du traitement disponible chez des adultes pour faire un essai randomisé de non-infériorité pour comparer deux traitements chez des enfants.

Cette étude illustre la construction d'un modèle basé sur l'*a priori* qui permettrait avec l'information disponible chez des adultes d'augmenter formellement les données concernant des enfants afin de minimiser l'expérimentation pédiatrique inutile. La discussion sur les lois *a priori* convenables pour tels designs est à la fois un élément nécessaire et souhaitable pour les essais bayésiens.

Donc en résumé, d'après les publications sur l'application des méthodes bayésiennes dans le cadre des maladies rares, nous recommandons l'approche bayésienne pour l'analyse des essais avec petite taille d'échantillons. Cette approche donne des résultats qui sont susceptibles d'aider les cliniciens à prendre une décision de traitement. L'inclusion de l'*a priori* est seulement acceptable si elle est basée sur des données "vraies", bien que l'utilisation des lois *a priori* non informative soit parfois préférée.

Avec les échantillons de petite taille, il est nécessaire d'estimer l'effet du traitement avec un modèle statistique afin d'ajuster les déséquilibres possibles sur des facteurs pronostiques entre traitements.

Chapitre 3

Classement des groupes de traitements et élicitation des avis des experts

3.1 Classement des groupes de traitements

Un problème fondamental dans la théorie de la décision est de savoir comment pondérer un ensemble d'études, de traitements ou de populations en fonction de leur importance. Cette importance est généralement jugée en fonction de plusieurs critères, tel que chaque critère peut être commun à un groupe ou bien à l'ensemble des groupes d'études.

Notre objectif est de pondérer les études, tel que les poids correspondent aux priorités des études. Le problème est alors de trouver la force ou les priorités de chaque étude par rapport à chaque critère puis de grouper les résultats obtenus pour chaque étude en tenant compte de chaque critère afin d'obtenir une seule priorité globale pour toutes les études.

Fréquemment, les critères eux-mêmes doivent être hiérarchisés ou classés en fonction d'un autre ensemble de critères (niveau supérieur). Les priorités ainsi obtenues sont ensuite utilisées comme facteur de pondération pour les études correspondantes aux critères.

Ce processus doit être poursuivi en comparant les études de niveau supérieur en termes de critères plus élevées et ainsi de suite jusqu'à arriver à un seul objectif global.

Dans la suite du chapitre, nous allons présenter des méthodes classiques de pondération pour classer les poids des critères dans chaque niveau hiérarchique en tenant compte des

critères du niveau supérieur.

3.2 Méthode du tri croisé

3.2.1 Introduction aux méthodes de tri croisé

La méthode de tri croisé introduit par [Thurstone, 1927] en 1927 était une étape importante dans la prise de décision scientifique. L'intérêt théorique et pratique de cette méthode est sa simplicité.

Dans le processus de prise de décision, de nombreux facteurs doivent être pris en considération simultanément et avec environ le même degré d'importance. Une approche avec plus de finesse est donc nécessaire pour obtenir une conclusion claire, d'où le recours au tri croisé.

Nous avons recours aux analyses croisées afin de décomposer le problème en plusieurs comparaisons d'éléments binaires, en croisant les éléments 2 à 2. Les comparaisons binaires à renseigner forment une matrice qu'on appelle matrice de comparaison binaire [Rahme et al., 2000].

Dans une matrice de comparaison binaire à n éléments de pondération, un expert ou bien un groupe d'experts indiquent si l'objet i est mieux qualifié que l'objet j ou inversement. Il est alors essentiel que les décideurs fassent plusieurs comparaisons par paires. En principe $n-1$ comparaisons devraient suffire, mais au moins $\frac{1}{2}n(n-1)$ comparaisons sont obligatoires (dans le cas d'une matrice réciproque). Les comparaisons d'un objet 1 à 2, ..., n , de l'objet 2 à 3, ..., n , et incluant $n-1$ à n ($\frac{1}{2}n(n-1)$) est le nombre de résultats pour les comparaisons 2 à 2 qui correspondent au nombre d'éléments d'une matrice $n \times n$, situés au-dessus ou en dessous de la diagonale.

| | Groupe 1 | Groupe 2 | Groupe 3 | Groupe 4 | Groupe 5 | Groupe 6 |
|----------|----------|----------|----------|----------|----------|----------|
| Groupe 1 | | | | | | |
| Groupe 2 | 1 vs 2 | | | | | |
| Groupe 3 | 1 vs 3 | 2 vs 3 | | | | |
| Groupe 4 | 1 vs 4 | 2 vs 4 | 3 vs 4 | | | |
| Groupe 5 | 1 vs 5 | 2 vs 5 | 3 vs 5 | 4 vs 5 | | |
| Groupe 6 | 1 vs 6 | 2 vs 6 | 3 vs 6 | 4 vs 6 | 5 vs 6 | |

Dans ce cas avec 6 groupes d'études il faut faire au moins *15 comparaisons*.

De ce fait, les experts fournissent les coefficients $a_{ij} > 0$ de la comparaison binaire, ces coefficients sont supposés être un remplaçant du coefficient w_i/w_j , w_i et w_j représentant respectivement les poids des groupes G_i et G_j .

Dans la suite du chapitre, nous définissons la matrice de comparaison binaire de dimen-

sion $n \times n$ comme une matrice carrée $A = [a_{ij}]$ avec $a_{ij} > 0$ pour tout $i, j = 1, \dots, n$.

Cette matrice est dite réciproque si $a_{ij} = \frac{1}{a_{ji}}$ pour tout $i, j = 1, \dots, n$ (donc automatiquement $a_{ii} = 1$ pour tout $i, j = 1, \dots, n$).

$$A = \begin{pmatrix} 1 & a_{12} & \cdots & a_{1n} \\ \frac{1}{a_{12}} & 1 & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{a_{1n}} & \frac{1}{a_{2n}} & \cdots & 1 \end{pmatrix}$$

La matrice des comparaisons ci-dessus est supposée être cohérente, dans le cas où toutes les comparaisons élémentaires sont cohérentes. La notion de cohérence, évoquée ici, est équivalente à la transitivité des comparaisons, c'est-à-dire :

$$a_{ij} \cdot a_{jk} = a_{ik} \quad \text{pour tout } i, j = 1, \dots, n.$$

Dans le cas où $a_{ij} = \frac{\omega_i}{\omega_j}$, avec $\omega_i \succ \omega_j$, alors que toute matrice cohérente est réciproque, une matrice réciproque n'est pas nécessairement cohérente.

Les matrices cohérentes correspondent à la situation idéale dans laquelle il y a des valeurs exactes des poids $\omega_1, \omega_2, \dots, \omega_n$. Les quotients $a_{ij} = \frac{\omega_i}{\omega_j}$ forment une matrice cohérente. Donc la matrice A peut s'écrire sous la forme :

$$A = \begin{pmatrix} 1 & \frac{\omega_1}{\omega_2} & \cdots & \frac{\omega_1}{\omega_n} \\ \frac{\omega_2}{\omega_1} & 1 & \cdots & \frac{\omega_2}{\omega_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\omega_n}{\omega_1} & \frac{\omega_n}{\omega_2} & \cdots & 1 \end{pmatrix}$$

Inversement, le point de départ de la théorie de comparaison appariée est le théorème de Saaty (voir [Saaty, 1977]) qui affirme que pour chaque matrice cohérente $A = [a_{ij}]$ de taille $n \times n$ il existe des nombres réels positifs $\omega_1, \omega_2, \dots, \omega_n$ pour tout $i, j = 1, \dots, n$.

Le défi de la méthode de comparaison appariée provient du manque de cohérence des matrices de comparaison qui se posent dans la pratique.

Étant donné une matrice A non cohérente de taille $n \times n$, la théorie tente de fournir une matrice cohérente B de taille $n \times n$ qui diffère de la matrice A aussi peu que possible. L'une des solutions possibles à ce problème a été proposée par Saaty ([Saaty, 1977]).

Il y a eu des discussions continues sur les méthodes qui permettent de trouver les meilleures solutions aux matrices de comparaisons. Parmi les compétiteurs les plus forts sont : la méthode principale aux valeurs propres, la méthode des moindres carrés, la régression logarithmique selon les moindres carrés ou les moyennes géométriques.

3.2.2 Méthode principale aux valeurs propres

La méthode la plus populaire pour estimer un vecteur prioritaire est celle proposée par Saaty lui-même, selon laquelle le vecteur prioritaire devrait être le vecteur propre principal de A . La solution dérivée de cette méthode est donnée en résolvant le problème des valeurs propres $A \omega = \lambda_{max} \omega$, or,

$$\sum_{j=1}^n a_{ij} \omega_j = \lambda_{max} \omega_i \quad i = 1, \dots, n$$

où $(\omega_1, \dots, \omega_n)$ sont les composantes du vecteur propre à droite associé à la plus grande valeur propre λ_{max} de A [Saaty and Vargas, 1984], [Brunelli, 2015], [Saaty and Hu, 1998]. λ_{max} est obtenu en résolvant le polynôme caractéristique

$$\det(A - \lambda_{max} I_n) = 0 \text{ avec } I_n \text{ la matrice identité.}$$

Les valeurs propres de A peuvent toutes être complexes à l'exception d'une. La valeur propre exceptionnelle dont l'existence est assurée par le théorème de Perron-Frobenius (Théorème 3.2.1) est notée λ_{max} et est réelle et positive. En addition, pour une matrice réciproque A , $\lambda_{max} \geq n$. Le vecteur propre correspondant ω est déterminé à un facteur multiplicatif près et ses composantes sont aussi réelles et positives. Par conséquent, la règle générale pour la résolution de notre problème pour une matrice réciproque positive arbitraire A est de déterminer la valeur caractéristique λ telle que $\lambda = \lambda_{max}$. C'est seulement pour cette valeur que l'on peut affirmer l'existence de l'échelle ratio ω [Saaty and Vargas, 1984], [Brunelli, 2015], [Saaty and Hu, 1998].

La solution de la méthode aux valeurs propres est obtenue itérativement comme :

$$\begin{cases} \omega = \lim_{k \rightarrow \infty} \frac{A^k e}{\|A^k\|} \\ \text{Avec } e = (1, \dots, 1)^T \\ \text{et } \|A^k\| = e^T A^k e \end{cases}$$

Théorème 3.2.1. *Théorème de Perron-Frobenius*

Soit A une matrice ≥ 0 irréductible. Alors

1. *A admet une valeur propre positive λ_{max} qui est la plus grande valeur propre (en module).*

2. Le vecteur propre de A associé à la valeur propre λ_{max} a des composantes positives et essentiellement uniques.
3. La valeur de λ_{max} est donnée par :

$$\lambda_{max} = \max_{x \geq 0} \min_{1 \leq i \leq n} ((Ax)_i / x_i) = \min_{x \geq 0} \max_{1 \leq i \leq n} ((Ax)_i / x_i)$$

Théorème 3.2.2. [Saaty, 1977]

Soit $A = [a_{ij}]$ une matrice $n \times n$ à coefficient positive $a_{ij} = \frac{1}{a_{ji}}$, alors A est cohérente si et seulement si $\lambda_{max} = n$

3.2.3 Méthode basée sur la régression logarithmique selon les moindres carrés ou Méthode des moyennes géométriques

Une deuxième méthode largement utilisée est la méthode basée sur la régression logarithmique selon les moindres carrés (R.L.M.C), proposée par Crawford [Crawford, 1987]. Selon cette méthode, chaque composante de ω est obtenue en tant que moyenne géométrique des éléments sur la ligne correspondante divisée par un terme de normalisation de sorte que $\sum_{i=1}^n \omega_i = 1$ [Brunelli, 2015],

$$u_i = \frac{\left(\prod_{j=1}^n a_{ij} \right)}{\sum_{i=1}^n \left(\prod_{j=1}^n a_{ij} \right)} \quad (3.1)$$

En analysant l'équation (3.1), il est évident que la méthode R.L.M.C est très intéressante pour les applications pratiques, car, contrairement à la méthode des vecteurs propres, les poids peuvent être exprimés comme des fonctions analytiques des entrées de la matrice, les coefficients des autres lignes ne contribuent pas directement au calcul des u_i . Ainsi, les relations incohérentes entre la ligne i et les autres rangées ne sont pas reflétées dans u_i . La méthode R.L.M.C est une procédure pour mettre en oeuvre l'idée que l'inverse d'une fonction d'un ensemble de variables, dont deux quelconques sont interchangeable, est égal à cette fonction appliquée à la réciproque des variables et donc la moyenne géométrique d'une ligne de A est l'inverse de la colonne correspondante de A . En outre, sur une note plus mathématique, il est intéressant de noter que le vecteur u obtenu avec cette méthode peut être obtenu de manière équivalente comme argument minimisant le problème d'optimisation suivant :

$$\left\{ \begin{array}{l} \text{minimiser} \\ \{u_1, \dots, u_n\} \end{array} \quad \sum_{i=1}^n \sum_{j=1}^n (\ln a_{ij} + \ln u_j - \ln u_i) \right. \quad (3.2)$$

$$\left. \begin{array}{l} \text{Avec} \\ \sum_{i=1}^n u_i = 1, \quad u_i > 0 \forall i \end{array} \right.$$

Ce problème d'optimisation comporte quelques interprétations. Nous savons que, dans le cas de la rationalité parfaite, $a_{ij} = u_i/u_j \forall i, j$.

En effet, on peut considérer $\sum_{i=1}^n \sum_{j=1}^n (a_{ij} - \frac{u_i}{u_j})^2$ comme distance entre la matrice A et la matrice $(\frac{u_i}{u_j})_{n \times n}$ associée au vecteur des poids u .

Une autre métrique peut être trouvée en utilisant le logarithme népérien \ln , qui est une fonction croissante monotone, obtenant ainsi, $\sum_{i=1}^n \sum_{j=1}^n (\ln a_{ij} + \ln(\frac{u_i}{u_j}))$. Le logarithme népérien d'un quotient est la différence des logarithmes népériens. Ensuite, le problème de minimisation (3.2) est introduit pour trouver un vecteur de priorité approprié associé à une "proche" approximation $(\frac{u_i}{u_j})_{n \times n}$ à la matrice A.

3.2.4 Méthode des moindres carrés(M.M.C)

Parmi toutes les méthodes étudiées dans la littérature, il apparaît que, en plus du développement de la méthode des valeurs propres et de la méthode basée sur la régression logarithmique selon les moindres carrés, la méthode des moindres carrés a été la source de développement.

En utilisant la méthode des moindres carrés le vecteur de priorité v est obtenu en résolvant le problème d'optimisation suivant :

$$\left\{ \begin{array}{l} \text{minimiser} \\ \{v_1, \dots, v_n\} \end{array} \quad \sum_{i=1}^n \sum_{j=1}^n (a_{ij} - \frac{v_i}{v_j})^2 \right. \quad (3.3)$$

$$\left. \begin{array}{l} \text{Avec} \\ \sum_{i=1}^n v_i = 1, \quad v_i > 0 \forall i \end{array} \right.$$

Ce problème d'optimisation (3.3) est difficile à résoudre, car la fonction est non linéaire et en général pas convexe, donc il n'existe ni une solution unique [Sándor, 2008] , ni une méthode numérique largement utilisée pour trouver la solution.

Soit A la matrice des comparaisons de taille $n \times n$,

$$A = \begin{pmatrix} 1 & a_{12} & \dots & a_{1n} \\ \frac{1}{a_{12}} & 1 & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{a_{1n}} & \frac{1}{a_{2n}} & \dots & 1 \end{pmatrix}$$

Le but est de trouver une matrice réciproque positive et cohérente X qui s'écrit sous la forme,

$$X = \begin{pmatrix} 1 & \frac{v_1}{v_2} & \dots & \frac{v_1}{v_n} \\ \frac{v_2}{v_1} & 1 & \dots & \frac{v_2}{v_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{v_n}{v_1} & \frac{v_n}{v_2} & \dots & 1 \end{pmatrix}$$

et qui minimise le carré de la norme de Frobenius

$$\|A - X\|_F^2 = \underset{\{v_1, \dots, v_n\}}{\text{minimiser}} \sum_{i=1}^n \sum_{j=1}^n \left(a_{ij} - \frac{v_i}{v_j} \right)^2$$

$$\text{Avec} \quad \sum_{i=1}^n v_i = 1, \quad v_i \succ 0 \quad \forall i \quad (3.4)$$

Des nouvelles variables x_1, x_2, \dots, x_{n-1} sont introduites comme suit :

$$x_1 = \frac{v_1}{v_2}, x_2 = \frac{v_1}{v_3}, \dots, x_i = \frac{v_1}{v_i}, \dots, x_{n-1} = \frac{v_1}{v_n} \quad (3.5)$$

Les formules inverses sont calculées à partir des équations (3.4) et (3.5) :

$$v_1 = \frac{1}{1 + \sum_{i=1}^{n-1} \frac{1}{x_j}}, \quad v_i = \frac{\frac{1}{x_{i-1}}}{1 + \sum_{i=1}^{n-1} \frac{1}{x_j}} \quad i = 2, 3, \dots, n \quad (3.6)$$

En se basant sur la formule (3.5) la matrice X peut s'écrire sous la forme

$$X = \begin{pmatrix} 1 & x_1 & x_2 & \dots & x_{n-1} \\ \frac{1}{x_1} & 1 & \frac{x_2}{x_1} & \dots & \frac{x_{n-1}}{x_1} \\ \frac{1}{x_2} & \frac{x_1}{x_2} & 1 & \dots & \frac{x_{n-1}}{x_2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{1}{x_{n-1}} & \frac{x_1}{x_{n-1}} & \frac{x_2}{x_{n-1}} & \dots & 1 \end{pmatrix}$$

et le problème d'optimisation de la méthode de moindres carrés (M.M.C) (3.3) est réduit à

$$\begin{aligned} \min & \quad f(x_1, x_2, \dots, x_{n-1}) \\ \text{Avec} & \quad x_1, x_2, \dots, x_{n-1} \succ 0 \end{aligned} \quad (3.7)$$

Où,

$$\begin{aligned}
 f(x_1, x_2, \dots, x_{n-1}) &= \|A - X\|_F^2 \\
 &= \sum_{j=2}^n \left[(a_{1j} - x_{j-1})^2 + \left(\frac{1}{a_{1j}} - \frac{1}{x_{j-1}} \right)^2 \right] \\
 &+ \sum_{i=2}^{n-1} \sum_{j=i+1}^n \left[(a_{ij} - \frac{x_{j-1}}{x_{i-1}})^2 + \left(\frac{1}{a_{ij}} - \frac{x_{i-1}}{x_{j-1}} \right)^2 \right]
 \end{aligned} \tag{3.8}$$

Étant donné que f est une fonction continue et dérivable pour $x_1, x_2, \dots, x_{n-1} \succ 0$, la condition nécessaire d'optimalité au premier ordre est la suivante [Sándor, 2008] :

$$\frac{\partial f}{\partial x_1} = \frac{\partial f}{\partial x_2} = \dots = \frac{\partial f}{\partial x_{n-1}} = 0 \tag{3.9}$$

Les dérivées partielles de premier ordre de f sont des fonctions rationnelles de x_1, x_2, \dots, x_n et peuvent être transformées en polynômes multivariés.

3.2.5 Cohérence

Un décideur parfaitement rationnel devrait être capable d'énoncer exactement ses préférences par paires, i.e $a_{ij} = \frac{\omega_i}{\omega_j} \quad \forall i, j$. Considérons donc les ramifications de cette condition sur les entrées de la matrice de comparaison par paires A .

Si nous écrivons $a_{ij}a_{jk}$ et appliquons la condition $a_{ij} = \frac{\omega_i}{\omega_j} \quad \forall i, j$ alors on peut dériver ce qui suit

$$a_{ij}a_{jk} = \frac{\omega_i}{\omega_j} \cdot \frac{\omega_j}{\omega_k} = \frac{\omega_i}{\omega_k} = a_{ik} \tag{3.10}$$

Alors, on remarque que, si toutes les entrées de la matrice de comparaison A satisfont à la condition $a_{ij} = \frac{\omega_i}{\omega_j}, \forall i, j$ Alors la condition suivante est vraie [Brunelli, 2015] :

$$a_{ik} = a_{ij}a_{jk} \quad \forall i, j, k \tag{3.11}$$

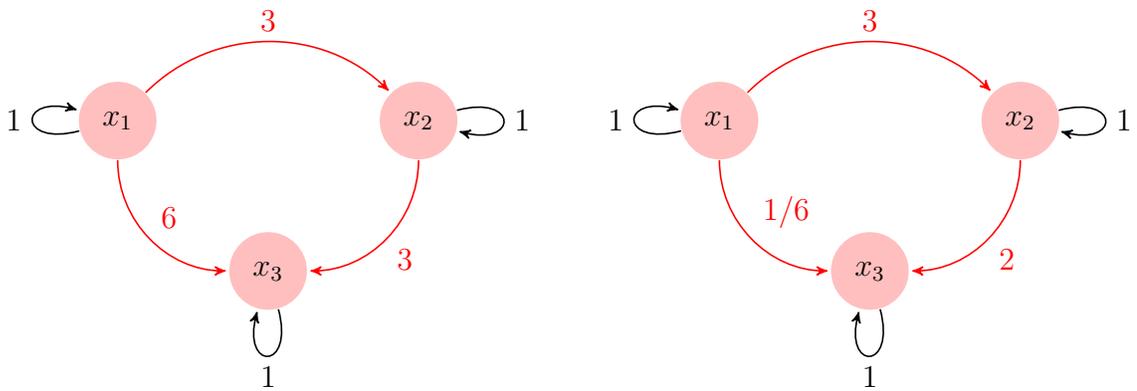
Ce qui signifie que chaque comparaison directe a_{ik} est exactement confirmée par toutes les comparaisons indirectes $a_{ij}a_{jk} \quad \forall j$.

Formellement, un décideur capable de donner des comparaisons par paires parfaitement cohérentes ne se contredit pas. Une matrice pour laquelle cette condition de transitivité est retenue est appelée cohérente [Brunelli, 2015].

Exemple 3.2.1. Nous considérons l'exemple de deux matrices de comparaisons par paires :

$$B = \begin{pmatrix} 1 & 3 & 6 \\ 1/3 & 1 & 2 \\ 1/6 & 1/2 & 1 \end{pmatrix} \quad C = \begin{pmatrix} 1 & 3 & 1/6 \\ 1/3 & 1 & 2 \\ 6 & 1/2 & 1 \end{pmatrix}$$

Pour lesquelles nous avons deux diagrammes présentés dans la figure 3.1 respectivement.



La matrice B est cohérente

La matrice C est non cohérente

Figure 3.1 – Exemple de transitivités cohérentes et non cohérentes

Être cohérent est rarement possible parce que de nombreux facteurs peuvent déterminer l'incohérence. Par exemple, le décideur pourrait être invité à utiliser des nombres entiers et leurs réciproques. Dans ce cas si $a_{ij} = 3$ et $a_{jk} = \frac{1}{2}$ il est impossible de trouver une valeur cohérente pour a_{ik} .

De plus, le nombre de transitivités indépendantes (i,j,k) dans une matrice d'ordre n est égale à $\binom{n}{3}$ mettant ainsi en évidence la difficulté d'être pleinement cohérent [Brunelli, 2015].

Exemple 3.2.2. Dans une matrice d'ordre 4, nous avons $\binom{4}{3} = 4$ transitivités indépendantes, qui sont (1, 2, 3), (1, 2, 4), (1, 3, 4), (2, 3, 4), c'est-à-dire le nombre d'affectations possibles des valeurs à i, j, k tel que $1 \leq i < j < k \leq 4$.

Malgré la difficulté d'être totalement transitif, il est indéniable que la cohérence est une propriété souhaitable. En fait, une matrice incohérente pourrait être un symptôme de l'incapacité ou de l'inexpérience du décideur dans le domaine.

En outre, il est possible d'envisager que les violations de la condition de cohérence (3.11) peuvent être de gravité différente et d'imaginer l'incohérence comme une notion progressive.

Par conséquent, au motif qu'une matrice doit s'écarter le moins possible de la condition de transitivité, un certain nombre d'indices d'incohérence ont été proposés dans la littérature pour quantifier cette déviation.

Il existe plusieurs indices d'incohérence dans la littérature et cette variété est en partie justifiée par le fait que la condition de cohérence peut être formulée de nombreuses façons équivalentes. Parmi eux, il faut compter les trois suivants :

- (i) $a_{ik} = a_{ij}a_{jk} \quad \forall i, j, k$
- (ii) Il existe un vecteur $(\omega_1, \dots, \omega_n)^T$ tel que $a_{ij} = \frac{\omega_i}{\omega_j} \quad \forall i, j$
- (iii) La matrice de comparaison appariée A a sa valeur propre maximale $\lambda_{max} = n$.

3.2.6 Mesure de la Cohérence

La méthode des valeurs propres a fourni un critère utile tel que la valeur propre maximale de la matrice de comparaison appariée A, λ_{max} , égale à n si et seulement si la matrice est cohérente ((Théorème 3.2.2)), Saaty propose l'index de cohérence [Saaty, 1977], [Saaty and Vargas, 1984]

$$\begin{aligned}
IC(A) &= \frac{\lambda_{max} - n}{n - 1} \\
&= -\frac{1}{n - 1} \sum_{i=2}^n \lambda_i
\end{aligned} \tag{3.12}$$

Preuve: A partir de $\text{trace}(A) = \sum_{i=1}^n \lambda_i = n$,

nous avons que $\lambda_{max} - n = -\sum_{i=2}^n \lambda_i$.

D'où l'égalité $\frac{\lambda_{max} - n}{n - 1} = -\frac{\sum_{i=2}^n \lambda_i}{n - 1}$. □

Par conséquent, plus la valeur de l'IC est importante, moins une matrice de comparaison deux à deux est cohérente.

Cependant, des études numériques ont montré que la valeur attendue de l'IC d'une matrice aléatoire de taille $n + 1$ est, en moyenne, supérieure à la valeur attendue de l'IC d'une matrice aléatoire d'ordre n . Par conséquent, il n'est pas exact de comparer l'IC des matrices de différents ordres et donc l'IC doit être rééchelonné.

Le ratio de cohérence, RC, est la version rééchelonnée de l'IC. Étant donné une matrice d'ordre n , le RC peut être obtenu en divisant l'IC par un nombre réel I_n (index aléatoire) qui n'est rien d'autre qu'une estimation de la moyenne de l'IC obtenue à partir d'un ensemble suffisamment grand de matrices de taille n générées de façon aléatoire [Koczkodaj, 1993], [Brunelli, 2015], alors

$$RC(A) = \frac{IC(A)}{I_n} \tag{3.13}$$

A noter que la génération de matrices aléatoires nécessite la définition d'une échelle limitée, où les entrées prennent des valeurs, par exemple l'intervalle $[1/9, 9]$. Selon Saaty [Saaty, 1987], une matrice de comparaison peut être considérée comme cohérente si sa valeur RC est inférieure à 0,1 (ou dans certains cas, 0,15) [Taji and Matsumoto, 2006] et rejeter pour une valeur RC supérieur à 0,1. Une valeur de RC = 0,1 signifie que les jugements sont à 10% aussi incohérents que s'ils avaient été donnés au hasard.

3.2.6.1 Choisir les différentes échelles

Dans la vie quotidienne, les gens sont plus enclins à utiliser des expressions linguistiques comme "je préfère légèrement" ou "je préfère fortement". Pour aider le décideur,

| n | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|----------|---|---|------|------|------|------|------|------|------|--------|
| I_n | 0 | 0 | 0.58 | 0.90 | 1.12 | 1.24 | 1.32 | 1.41 | 1.45 | 1.4854 |

Table 3.1 – Les valeurs de l'index I_n

certaines expressions linguistiques ont été proposées et ensuite liées à différentes valeurs affectables aux entrées a_{ij} . Par conséquent, le décideur peut exprimer des opinions sur les couples en utilisant des termes linguistiques, qui sont alors associés aux nombres réels.

Dans son article sur les processus analytiques hiérarchiques (AHP) [Saaty, 1987], Saaty a proposé une association entre les jugements verbaux et les valeurs pour les comparaisons par paires. D'autres échelles ont été proposées et étudiées, l'une des plus importantes est l'échelle équilibrée proposée par Pöyhönen [Pöyhönen et al., 1997].

L'échelle équilibrée, et l'échelle de Saaty ainsi que leurs correspondances avec les jugements verbaux sont rapportées dans la table 3.2 .

Table 3.2 – Les échelles fondamentales

| Description verbale | Explication | Échelle de Satty | Échelle équilibrée |
|---|---|------------------|--------------------|
| Indifférence | Deux activités contribuent également à l'objectif | 1 | 1 |
| Des valeurs intermédiaires entre les deux jugements adjacents | Quand un compromis est nécessaire | 2 | 1.22 |
| Importance modérée | L'expérience et le jugement favorisent fortement une activité par rapport à une autre | 3 | 1.5 |
| Des valeurs intermédiaires entre les deux jugements adjacents | Quand un compromis est nécessaire | 4 | 1.86 |
| Forte Préférence | L'expérience et le jugement favorisent très fortement une activité par rapport à une autre | 5 | 2.33 |
| Des valeurs intermédiaires entre les deux jugements adjacents | Quand un compromis est nécessaire | 6 | 3 |
| Préférence très forte ou démontrée | Une activité est fortement favorisée et sa supériorité est démontrée dans la pratique | 7 | 4 |
| Des valeurs intermédiaires entre les deux jugements adjacents | Quand un compromis est nécessaire | 8 | 5.67 |
| Préférence extrême | Les preuves favorisant une activité par rapport à une autre donnent une affirmation de l'ordre le plus élevé possible | 9 | 9 |

Pour savoir quelle échelle est meilleure un débat est ouvert, quoique l'échelle de Saaty ne semble pas optimale [Brunelli, 2015]. Il est généralement admis, alors que d'autres échelles semblent être plus acceptées. Par exemple, l'échelle équilibrée a été proposée sur la base d'expériences empiriques sur des personnes. Logiquement, cette échelle nécessitera plus de recherches du point de vue comportemental que du point de vue mathématique [Ishizaka and Labib, 2009].

Enfin, malgré le débat ouvert sur l'association entre les étiquettes linguistiques et les valeurs numériques, il y a un accord sur l'utilisation d'échelles numériques bornées, dont le plus célèbre est l'ensemble de tous les nombres entiers jusqu'à 9 et leurs réciproques :

$$\left\{ \frac{1}{9}, \dots, \frac{1}{2}, 1, 2, \dots, 9 \right\}$$

3.2.6.2 Exemple de calcul de l'Index de Cohérence

Supposons une matrice de comparaison appariée telle que représentée dans le tableau 3.3.

| | F_1 | F_2 | F_2 | F_4 |
|---------------------------|-------------|-------------|-------------|--------------|
| F_1 | 1 | 1/3 | 1/2 | 9 |
| F_2 | 3 | 1 | 5/2 | 1/7 |
| F_3 | 2 | 2/5 | 1 | 3/2 |
| F_4 | 1/9 | 7 | 2/3 | 1 |
| Somme des colonnes | 6.12 | 8.73 | 4.67 | 11.64 |

Table 3.3 – Matrice de comparaison par paires

Ensuite, nous normalisons la somme des colonnes en utilisant la formule suivante :

$$a_{ij}' = \frac{a_{ij}}{\sum_{i=1}^n a_{ij}}$$

Nous obtenons alors la matrice normalisée suivante (Tableau 3.4)

L'étape suivante est de calculer λ_{max} , on utilise la formule :

$$AX = \lambda_{max} X \tag{3.14}$$

| | F_1 | F_2 | F_2 | F_4 |
|--------------|----------|----------|----------|----------|
| F_1 | 0.16 | 0.04 | 0.11 | 0.77 |
| F_2 | 0.49 | 0.11 | 0.54 | 0.01 |
| F_3 | 0.33 | 0.05 | 0.21 | 0.13 |
| F_4 | 0.02 | 0.8 | 0.14 | 0.09 |
| Somme | 1 | 1 | 1 | 1 |

Table 3.4 – Matrice normalisée

tel que AX est obtenu en multipliant la matrice du tableau 3.3 avec la première colonne de la matrice normalisée du tableau 3.4 :

| | A | | | | * | X | = | AX |
|-------|----------|-------|-------|-------|---|----------|---|-----------|
| | F_1 | F_2 | F_2 | F_4 | | | | |
| F_1 | 1 | 1/3 | 1/2 | 9 | | 0.16 | | 0.67 |
| F_2 | 3 | 1 | 5/2 | 1/7 | | 0.49 | | 1.80 |
| F_3 | 2 | 2/5 | 1 | 3/2 | | 0.33 | | 0.90 |
| F_4 | 1/9 | 7 | 2/3 | 1 | | 0.02 | | 3.70 |

Table 3.5 – Calcul de la valeur propre maximale

Donc pour trouver λ_{max} il suffit de résoudre l'équation 3.14 : $\lambda_{max} = \text{moyenne} \left\{ \frac{AX}{X} \right\}$

Alors, $\lambda_{max} = \text{moyenne} \left\{ \frac{0.67}{0.16}, \frac{1.80}{0.49}, \frac{0.90}{0.33}, \frac{3.70}{0.02} \right\} = 48.9$

En utilisant la formule 4.3, nous calculons l'index de cohérence :

$$IC = \frac{\lambda_{max} - n}{n - 1} = \frac{48.9 - 4}{4 - 1} = 15$$

Maintenant nous pouvons calculer le ratio de cohérence (RC) en utilisant la formule 3.13 qui consiste à diviser l'IC par l'index aléatoire I_n Tableau 3.1, $CR = 15/0.90 = 16.67$. Selon Saaty [Saaty, 1987], la matrice de comparaison ne peut pas être considérée comme cohérente, car sa valeur RC est supérieure à 0,1.

3.3 Utilisation des avis d'experts pour trouver l'*a priori* dans les modèles statistiques bayésiens

L'information *a priori* dans un modèle bayésien peut être obtenue à partir des études antérieures, ou bien en leur absence, à partir des connaissances d'experts. Cette utilisation du cadre bayésien reflète les *a priori* ou bien les estimations initiales lorsqu'ils sont mis à jour lorsque de nouvelles données deviennent disponibles.

Dans cette section, nous présentons un cadre pour un design statistique du processus d'élicitation des avis experts pour quantifier les connaissances d'experts sous une forme appropriée pour la saisie comme information *a priori* dans les modèles bayésiens.

Nous identifions cinq éléments incontournables : déterminer le but et la motivation de l'utilisation de l'information *a priori*, spécifier les connaissances pertinentes disponibles des experts, la formulation du modèle statistique, la gestion de l'incertitude, et la conception d'un protocole pratique d'élicitation.

Dans la modélisation statistique bayésienne, l'"élicitation des experts" se réfère au processus d'obtention d'avis d'experts, ainsi que l'incertitude, qui sont ensuite formulés avec circonspection dans les distributions informatives *a priori* [Choy et al., 2009].

3.3.1 Questionnaires et élicitation de l'opinion des experts

L'élicitation de l'opinion des experts est un processus formel qui consiste à obtenir d'informations ou de réponses à des questions précises sur certains enjeux qui sont nécessaires pour atteindre certains objectifs analytiques. Nous allons présenter dans cette section les principales étapes requises par le statisticien, en particulier les efforts initiaux nécessaires à la conception de l'élicitation [Choy et al., 2009].

Plusieurs étapes sont nécessaires :

- Déterminer le but et la motivation de l'utilisation de l'information *a priori*.
- Spécifier une connaissance *a priori* disponible auprès d'experts ou d'autres sources, afin de définir un objectif approprié et réalisable de l'élicitation.
- Formuler un modèle statistique qui représente le modèle conceptuel. Définir la vraisemblance $p(y|\theta)$ caractérisant le modèle de données et la loi *a priori* $p(\theta)$ reflétant les connaissances *a priori* disponibles.
- Concevoir le codage numérique pour l'élicitation efficace de l'information *a priori* et la représentation de la distribution statistique.

→ Prendre une décision cruciale : si une approche directe ou indirecte est plus ap-

propriété à l'élicitation .

→ Sélectionnez les statistiques devant être prouvées, leur ordre et les unités appropriées.

→ Déterminer la méthode de communication appropriée pour obtenir les informations nécessaires (visuelle, sous forme de tableau, verbale y compris la formulation des questions, et ainsi de suite).

→ Spécifier la méthode d'estimation pour calculer les paramètres du modèle *a priori* des renseignements élicités, ceci est particulièrement important pour les méthodes indirectes

→ Déterminer le poids relatif de la loi *a priori* et de la vraisemblance, par exemple, en spécifiant *a priori* la taille effective de l'échantillon ou une autre fonction de la variance *a priori*.

■ Gérer l'incertitude pour l'élicitation précise et robuste.

→ Considérer l'élicitation de multiples experts

→ Conditionner les experts à d'éventuels biais

→ Concevoir un modèle d'élicitation pour minimiser les biais importants et vérifier les informations élicitées.

→ Effectuer une analyse de sensibilité aux hypothèses *a priori* relatives à l'élicitation.

→ Valider les résultats élicités, par exemple, via la transmission des commentaires ou de l'étalonnage.

■ Concevoir un protocole d'élicitation pour gérer la logistique de la mise en oeuvre de l'élicitation.

→ Sélectionner les contributions et les autres sources d'information *a priori*.

→ Déterminer la contribution relative des experts et d'autres sources, par exemple, en mettant en commun (pooling) ou bien en utilisant des méthodes plus sophistiquées telles que la méta-analyse.

→ Choisir différents moyens de diffusion tels que des questionnaires, des entretiens individuels, etc.

→ Concevoir soigneusement le design de la préparation des experts, y compris la motivation et la formation.

Cependant, parfois, un processus formel est impossible en raison d'une variété de raisons, telles que la logistique de la convocation d'une réunion de tous les experts en même temps

[Ayyub, 2000].

Nous suggérons une procédure pour effectuer l'élicitation de l'opinion des experts grâce à des questionnaires. La principale difficulté dans la conception de questionnaires aux fins de l'élicitation de l'opinion des experts est que leur conception doit assurer les conditions suivantes :

1. Communiquer correctement l'intérêt des questions aux experts.
2. Éliminer toute ambiguïté ou imprécision dans les déclarations des questions et les réponses attendues.
3. Éliminer toute ambiguïté ou imprécision dans la façon dont les réponses doivent être exprimées.
4. Fournir un design efficace complet, clair et facile à suivre.

Une approche similaire au processus formel d'élicitation de l'opinion des experts peut être utilisée pour la construction et l'administration des questionnaires. Les étapes nécessaires sont semblables au processus formel d'élicitation de l'opinion des experts présenté dans le paragraphe précédent avec une différence principale concernant l'élaboration et le test des questionnaires. Cette étape doit être effectuée par une équipe d'analyse des risques. Pour les questions choisies, les questionnaires doivent être conçus de telle sorte que chaque question aborde séparément un problème spécifique. Cependant, des questions similaires peuvent être traitées par la même structure de questionnaire avec quelques changements dans son contenu [Ayyub, 2000]. Une conception du questionnaire devrait inclure les éléments suivants :

1. Description du problème.
2. Familiarisation des experts avec la problématique.
3. Aspects de la question qui devraient être pris en compte dans son évaluation.
4. Aspects de la question qui ne devraient pas être pris en compte dans son évaluation.
5. Contenu, unités et nature des réponses attendues.

Les questionnaires développés doivent être testés avant leur utilisation dans le processus d'élicitation des opinions des experts. Le groupe des experts peut être sélectionné sur les bases de leur familiarisation avec les questions, les objectifs de l'étude, la disponibilité et la volonté de fournir des réponses opportunes.

3.4 Prise de décision en présence de plusieurs experts

Les chercheurs font souvent appel à plus d'un expert quand il demeure beaucoup d'incertitude au moment de la prise de décision. Il peut alors être nécessaire ou souhaitable d'obtenir une distribution unique qui englobe les croyances de plusieurs experts. Comment peut-on donc réunir les informations de ces experts? Nous allons présenter deux approches pour combiner les différents *a priori* en un seul consensus d'*a priori* [O'Hagan et al., 2006]. Ces deux approches sont l'agrégation mathématique ou comportementale au sein de chaque approche globale, il existe un certain nombre de méthodes et de variations spécifiques [O'Hagan et al., 2006].

3.4.1 Agrégation mathématique

Considérons la situation où les différentes lois de probabilité ont été élicitées par un certain nombre d'experts. Nous souhaitons combiner ces lois pour produire une seule loi d'agrégation. Alors que les lois individuelles des experts sont clairement censées représenter les points de vue réels et les connaissances des experts correspondants, la loi d'agrégation ne peut pas avoir la même interprétation. Nous allons présenter ensuite deux méthodes en vigueur de l'agrégation mathématique qui sont la méthode par moyenne des opinions (averaging opinions) et l'approche par mélange des opinions (pooling opinions).

3.4.1.1 Méthode par moyenne des opinions

La méthode proposée par [Cooke, 1991] est un modèle de calcul de la moyenne pondérée basée sur la performance en utilisant les propriétés des règles de notation (scoring rules). Il souligne que l'objectif fondamental de la science est de construire un consensus rationnel et, par conséquent, le processus de collecte des avis d'experts doit être soumis aux cinq principes de base suivants [Fumika, 2004] :

1. Tous les résultats doivent être reproductibles, avec des modèles de calcul et des données clairement spécifiées et mis à disposition.
2. La source des données (nom et institution) doit être identifiée, et les données doivent correspondre à la source exacte à partir de laquelle les données sont "élicitées".
3. Les évaluations d'experts doivent être, en principe, physiquement observables.
4. Le processus d'élicitation doit veiller à ce que les opinions réelles des experts soient recueillies.

5. Tous les experts doivent être considérés de manière équivalente avant le processus d'agrégation.

Le modèle classique est conçu pour satisfaire tous ces principes afin de trouver un consensus rationnel.

En cas de variables continues, le modèle exige des experts la fourniture d'un ensemble de quantiles fixes $q_r = \{q_1, \dots, q_R\}$, pour certaines variables inconnues, X_1, \dots, X_n . Les décideurs déterminent alors l'intervalle intrinsèque $[q_l, q_h]$ de chaque variable pour chaque expert, tel que q_l et q_h correspondent respectivement à la limite inférieure et supérieure des quantiles. Les poids pour les experts sont établis par deux mesures de performance : la calibration et les scores d'information. La calibration $C(e)$ d'un expert e , est la probabilité statistique qu'un quantile fournit par un expert corresponde à un ensemble de résultats expérimentaux [O'Hagan et al., 2006], [Fumika, 2004].

Notations :

- Soit $p = \{p_1, \dots, p_n\}$ la distribution de probabilité d'une variable aléatoire discrète prenant n valeurs possibles.
- Soit $s = \{s_1, \dots, s_n\}$ la distribution empirique à partir de N échantillons indépendants de la distribution p (c.-à-d. N est le nombre de variables).
- La différence entre p et s peut être mesurée par les informations relatives de s en tenant compte de p . Elle est basée sur la distance de Kullback-Leiber (K-L) entre deux probabilités de distributions discrètes. Donc, la distance K-L entre p et s est donnée par :

$$I(s, p) = \sum_{i=1}^n s_i \ln(p_i/s_i)$$

- [Cooke, 1991] note que pour N variables $P(2NI(s, p) \leq x) \rightarrow \chi_{n-1}^2(x)$ si l'expert est parfaitement calibré.
- $C(e)$ est la probabilité qu'une telle variable aléatoire Chi-deux devrait dépasser la valeur observée de $2N I(s, p)$, d'où :

$$C(e) = 1 - \chi_{n-1}^2(2NI(s, p))$$

Le score de l'information mesure le degré d'intensité de la distribution d'un expert par rapport à des mesures déterminées, généralement d'une distribution uniforme ou bien log-uniforme sur un intervalle intrinsèque pour chaque variable.

L'information relative pour chaque expert e sur une variable est obtenue sous forme :

$$I(e) = \sum_{i=1}^n p_i \ln(p_i/r_i) \quad \text{avec } r_i \text{ mesure d'information pour l'intervalle } i$$

Le score global de l'information de chaque expert est la moyenne des scores de l'information sur toutes les variables.

Les poids pour les experts qui répondent aux règles de notation asymptotiquement appropriées sont ensuite déterminés comme :

$$w_e = \frac{w'_e}{\sum w'_e} \quad \text{avec } w'_e = C(e) \times I(e) \times I_\alpha(C(e)).$$

Notons pour un seuil de signification α , tel que $0 < \alpha < 1$

$$\begin{cases} I_\alpha(C(e))=1 & \text{Pour } C(e) \geq \alpha \\ I_\alpha(C(e))=0 & \text{Pour } C(e) < \alpha \end{cases}$$

Étant donné un poids w_e et une fonction de répartition F_e pour chaque expert, la distribution combinée des experts est donnée par : $\sum_e w_e F_e$.

3.4.1.2 Méthode par mélange des opinions

La méthode par mélange des opinions des experts consiste à combiner les avis d'experts par un processus d'établissement des moyennes à chaque point de la fonction de répartition(CDF). Il existe deux types de méthodes couramment utilisées pour mélanger les opinions : linéaire et logarithmique.

Linear opinion pooling Une méthode mathématique couramment utilisée, en raison à la fois de sa facilité et sa simplicité, est une moyenne arithmétique pondérée des lois de probabilités individuelles. Sa fonction simple est donnée par : $p(\theta) = \sum_{i=1}^n \omega_i p_i(\theta)$.

Avec $p_i(\theta)$ la loi de probabilité d'un expert i pour une incertitude θ , $p(\theta)$ représente la loi de probabilité agrégée, ω_i correspond au poids donné par chaque expert, ce sont les valeurs non négatives et $\sum_{i=1}^n \omega_i = 1$.

Un inconvénient de cette approche est que lorsque de nouvelles données sont obtenues, on pourrait soit mettre à jour la distribution de chaque expert individuellement, puis combiner la distribution postérieure résultante en utilisant la méthode de mélange linéaire, ou bien de tout d'abord combiner les distributions des experts et ensuite mettre à jour la

distribution du consensus [O'Hagan et al., 2006] .

À l'exception de quelques cas triviaux, ces deux options ne donnent pas le même résultat. Par conséquent, un décideur pourrait être confronté à deux distributions *a posteriori* différentes impliquant deux décisions optimales différentes, mais sans aucun moyen justifiable de choisir entre eux. De même, lorsqu'on élicite une loi conjointe pour deux variables aléatoires X et Y, il se peut que tous les experts conviennent que X et Y sont indépendants, mais un mélange linéaire de leurs lois conjointes donne généralement une loi dans laquelle elles ne sont plus indépendantes [O'Hagan et al., 2006].

Logarithmic opinion pooling Une approche alternative, est la méthode de mélange des opinions logarithmique, dans lequel le consensus est obtenu en prenant une moyenne géométrique pondérée des n distributions individuelles,

$$p(\theta) = k \prod_{i=1}^n p_i(\theta)^{\omega_i}$$

Où k est une constante de normalisation, l'interprétation des poids se fait comme dans l'approche précédente, et si le décideur est également confiant dans les capacités des experts, il est naturel de choisir $\omega_i = \frac{1}{n} \forall i$ [O'Hagan et al., 2006].

Le mélange logarithmique des opinions est "externally bayesian" lorsque les poids $\sum_{i=1}^n \omega_i = 1$, et est également compatible avec les jugements d'indépendance, mais expose à la marginalisation.

Supposons que nous ayons deux événements mutuellement exclusifs, A et B, et définissons C comme l'événement «A ou B», de sorte que pour tout individu cohérent $P(C) = P(A) + P(B)$. Supposons maintenant que chaque expert indique ses propres probabilités pour les événements A et B. Pour obtenir une probabilité regroupée P(C), on pourrait soit calculer P(C) pour chaque expert et regrouper les résultats, soit regrouper d'abord les probabilités pour A et B séparément, puis additionner les résultats.

Dans le cadre de la méthode de mélange logarithmique, ces deux options ne donnent pas le même résultat sauf dans quelques cas triviaux.

3.5 Agrégation comportementale

Les agrégations comportementales ont été largement étudiées dans le domaine de la prise de décision de groupe et du comportement organisationnel. L'hypothèse importante

des approches comportementales est que, par l'échange d'opinions ou d'informations, les experts peuvent éventuellement atteindre un équilibre où une interaction ultérieure ne changera pas leurs opinions. L'un des approches comportementales les plus connues est la technique Delphi.

L'approche générale de la méthode de Delphi est applicable à divers types de jugement, de groupe et de prise de décision. Cependant, dans la situation où il s'agit d'éliciter une loi de probabilité de consensus de groupe, chaque individu fournit d'abord sa propre distribution avec une explication de ses vues. Cette distribution est ensuite fournie à tous les autres experts du groupe. L'expert a alors la possibilité de réviser ses propres croyances à la lumière des opinions exprimées par d'autres experts.

Ce processus peut être itéré plusieurs fois, jusqu'à ce que les experts convergent vers une seule distribution ou que les facilitateurs décident d'agréger mathématiquement les différentes distributions.

3.6 Conclusion

Le débat sur l'agrégation mathématique versus l'agrégation comportementale n'est pas encore résolu, car il n'y a pas de raison de favoriser une méthode plutôt que l'autre et chaque approche a ses avantages et ses inconvénients. Les approches mathématiques sont faciles à utiliser. Mais elles n'ont pas de normes pour choisir des poids. En outre, plusieurs résultats montrent qu'aucune fonction d'agrégation ne peut satisfaire toutes les propriétés souhaitées d'un pool d'opinions, à moins que l'opinion commune ne converge vers une seule opinion individuelle, ce qui implique dans ce cas la présence d'un leader d'opinion.

Les approches comportementales permettent aux experts d'améliorer dynamiquement leurs informations et de réviser leurs opinions au cours des interactions, mais pour certains d'entre eux elles ne sont pas fixées ou complètement spécifiées et ne peuvent donc garantir la convergence ou la répétabilité.

Chapitre 4

Étude de l'efficacité du méthotrexate associé aux corticostéroïdes par rapport à la corticothérapie seule : Résultats des méthodes bayésiennes appliquées

4.1 Introduction

La conception et l'interprétation des essais cliniques comparant des traitements dans le cas des études de maladies rares sont limitées par la disponibilité d'échantillons de petites tailles, même lorsque les patients sont recrutés sur plusieurs années. Pour ces études, la taille de l'échantillon nécessaire pour répondre à une exigence de puissance fréquentiste est clairement une limite. C'est la raison pour laquelle nous avons besoin de nouvelles méthodes et théories pour relever les défis de l'organisation de l'évaluation thérapeutique dans le cadre des maladies rares.

Nous proposons dans ce chapitre une approche bayésienne non utilisée jusque là dans la littérature pour la conduite des essais de maladies rares comparant un traitement expérimental avec un contrôle. Une élicitation par des cliniciens de leurs croyances concernant l'efficacité d'un nouveau traitement est utilisée pour le choix de l'*a priori* bayésien pour estimer les paramètres inconnus du modèle.

La détermination de la loi *a priori* dans le modèle bayésien est difficile. Nous allons présenter ci-dessous d'une part les méthodes qui permettent de déterminer la loi *a priori* en utilisant l'avis des experts et d'autre part une méthode permettant de considérer des informations provenant d'études historiques et/ou de données provenant d'autres études "voisines".

D'une part, nous proposons une approche bayésienne [Hampson et al., 2014] pour la conduite d'essais de maladies rares comparant un traitement expérimental (méthotrexate associé aux corticostéroïdes) à un traitement de référence (corticostéroïdes seuls) où les réponses des patients sont classées comme succès ou échec.

D'autre part, nous utilisons la méthode de l' ϵ -contamination, qui se base sur le principe de contaminer une loi *a priori*, pas entièrement satisfaisante, par une série de lois provenant des informations d'autres études portant sur la même maladie, le même traitement ou concernant la même population.

4.2 Présentation de l'artérite à cellules géantes

L'artérite à cellules géantes (GCA) ou maladie de Horton est une forme de vascularite, un groupe de troubles associés à une inflammation de la paroi de certaines artères. Elle touche les artères de grand ou de moyen calibre. Les artères du crâne, le plus souvent les branches de l'artère carotide externe, sont les plus impliquées, surtout les artères temporales, mais d'autres artères peuvent également être touchées. L'inflammation provoque un rétrécissement de la lumière artérielle, ce qui entraîne des troubles circulatoires. Les symptômes peuvent inclure un mal de tête lancinant sur un côté ou à l'arrière de la tête, une fièvre $> 38^{\circ}C$, une douleur à la mastication, et/ou des problèmes visuels; le risque majeur de cette affection est une perte de la vision. La cause de la GCA n'est pas élucidée, mais une réponse immunitaire anormale a été constatée. Plusieurs facteurs génétiques et environnementaux peuvent accroître le risque de développer une GCA. Un traitement précoce est essentiel, fondé essentiellement sur la corticothérapie qui est le traitement de référence. La corticothérapie doit être prolongée au moins 18 mois avec une défervescence très progressive, car le risque de rechute est important. Une corticodépendance peut se développer. Elle se caractérise par la réapparition des symptômes lors de la diminution de la posologie des corticoïdes. Le recours à d'autres immunosuppresseurs (le méthotrexate surtout, l'azathioprine, ou le cyclophosphamide) ou aux antipaludéens de synthèse peut alors être indiqué.

Les corticoïdes sont très efficaces. Toutefois, leur administration prolongée expose à des effets secondaires fréquents. Une surveillance étroite est indispensable et le traitement ne

doit jamais être interrompu subitement. Les principaux effets secondaires sont, l'hypertension artérielle, des troubles de sommeil et de l'humeur, des troubles hormonaux, une perte de masse musculaire, des troubles digestifs, une prise de poids et un risque accru d'infection.

L'artérite à cellules géantes (GCA) est la vascularite la plus fréquente chez les personnes âgées de plus de 50 ans (âge médian au diagnostic entre 70-75 ans) avec une incidence annuelle entre 1 cas sur 3000 et 1 cas sur 25000. La maladie de Horton survient deux fois plus fréquemment chez les femmes que chez les hommes. D'autres facteurs augmentent le risque de développer une GCA comme être de descendance d'Europe du Nord (en particulier scandinave), les personnes nées dans les pays d'Europe du Nord semblent avoir des taux plus élevés de GCA et les patients ayant une pseudo-polyarthrite rhizomélisque - environ 15% des personnes atteintes de pseudo-polyarthrite rhizomélisque ont également une GCA.

La corticothérapie est le traitement de référence de la maladie de Horton. Cependant, elle est souvent compliquée d'effets secondaires plus ou moins sévères, en raison d'une posologie initiale conséquente et d'une dose cumulée souvent importante, chez des patients âgés. En adjonction des corticoïdes la méthotrexate est l'un des traitements alternatifs les plus utilisés dans les cas où la corticothérapie est en défaut (effets secondaires importants, cortico dépendance,..) malgré son faible niveau de preuve et sa faible efficacité. Le cyclophosphamide et le tocilizumab semblent prometteurs, mais nécessitent une validation préalable. Les résultats des anti-TNF- α et de l'azathioprine sont décevants.

4.3 Première Méthode : Modèle bayésien pour la conduite d'un essai de non-infériorité dans le cas des maladies rares

4.3.1 Introduction de la méthode

Lorsque des données historiques pertinentes ne sont pas disponibles, les distributions *a priori* doivent être déterminées à partir d'un avis d'expert. Une littérature riche a été écrite sur les méthodes d'élicitation (Chapitre 3).

Dans les applications d'essais cliniques, le recueil d'une opinion préalable a été mis en oeuvre par divers moyens, y compris les réunions en face à face [Johnson et al., 2011], le courrier électronique, le téléphone et les questionnaires postaux [Tan et al., 2008], [Chaloner and Rhame, 2001].

Les réunions en face à face ont l'avantage de permettre de former et de pouvoir fournir des commentaires aux experts, ce qui augmente les chances que la distribution de probabilité obtenue reflète une représentation adaptée de leurs croyances.

Nous utilisons un schéma bayésien proposé par [Hampson et al., 2014] pour la recherche clinique dans les maladies rares.

Une analyse bayésienne des petits échantillons peut être conduite dans le but d'améliorer les connaissances dans la compréhension du traitement, mais ne peut pas montrer de manière sûre une différence entre deux traitements. Une prise de décision dans ce contexte repose donc sur un faisceau d'arguments.

Les méthodes utilisées ici doivent être considérées comme un dernier recours quand une taille d'échantillon satisfaisante ne peut être obtenue dans des temps raisonnables et ne doivent être envisagées que quand un essai classique ne peut être mis en place dans les règles de l'art.

Notre application clinique concerne un essai randomisé multicentrique de non-infériorité comparant le méthotrexate (MTX) associé à la corticothérapie à la corticothérapie (CS) seule pour le traitement de l'artérite à cellules géantes. La marge de non-infériorité est pré spécifiée par les cliniciens. Trois experts français dans le domaine ont participé au choix des lois *a priori* à partir des données existantes et de leur expertise.

4.3.2 Modèle bayésien

On suppose qu'on a n patients recrutés, tels que n_1 est le nombre de patients qui ont reçu le traitement classique CS et n_2 le nombre de patients qui ont reçu le traitement MTX. Lorsque l'essai est complet, on aura S nombre de succès et F nombre d'échecs, tel que S_j et F_j correspondent successivement au succès (qui est le nombre de patients n'ayant aucune rechute) et à l'échec du traitement (les patients ayant une ou plusieurs rechutes) $j; j=1,2$.

On note P_1 et P_2 la probabilité de la réussite du traitement dans les groupes CS et MTX respectivement.

La loi *a priori* de P_1 , va être exprimée comme une distribution bêta, avec les paramètres a et b ($P_1 \approx \text{Beta}(a,b)$, car $P_1 \in [0, 1]$)

La loi *a priori* de P_2 (v.a. du taux de succès du traitement MTX) va être exprimée comme une distribution bêta, avec les paramètres c et d ($P_2 \approx \text{Beta}(c,d)$, car $P_2 \in [0, 1]$)

Donc P_1 et P_2 sont supposés indépendantes.

On mesure l'avantage de MTX par rapport au CS en utilisant le log OR :

$$\theta = \log \frac{P_2/1 - P_2}{P_1/1 - P_1} = \log \frac{P_2(1 - P_1)}{P_1(1 - P_2)}$$

La loi jointe *a priori* sera complétée en précisant la loi de θ que l'on choisit, indépendante de celle de P_1 .

La loi de θ est choisie comme une loi normale indépendante de la loi *a priori* de $P_1 \approx \text{Beta}(a,b)$.

Donc,

$$\theta \approx N [\mu, \sigma^2]$$

La distribution jointe de P_1 et θ est :

$$f_0(P_1, \theta) = h_0(P_1)k_0(\theta)$$

$$\text{avec : } \begin{cases} h_0(P_1) &= \frac{1}{\beta(a,b)} P_1^{a-1} (1 - P_1)^{b-1} \\ k_0(\theta) &= \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2} (\theta - \mu)^2\right) \end{cases}$$

La loi jointe *a priori* correspondante pour P_1 et P_2 est :

$$g_0(P_1, P_2) \propto \frac{P_1^{a-1}(1-P_1)^{b-1}}{P_2(1-P_2)} \cdot \exp \left[-\frac{1}{2\sigma^2} \cdot \left(\log \frac{P_2(1-P_1)}{P_1(1-P_2)} - \mu \right)^2 \right] \quad (4.1)$$

Donc, la loi jointe de P_1 et θ est présentée comme suit :

$$\begin{aligned} f_0(P_1, \theta) &= f_0(P_1, h(P_1, P_2)) \\ &= h_0(P_1)k_0(h(P_1, P_2)) \\ &\propto P_1^{a-1}(1-P_1)^{b-1} \frac{1}{\sqrt{2\pi}\sigma} \exp \left(-\frac{1}{2\sigma^2} \cdot \left[\log \frac{P_2(1-P_1)}{P_1(1-P_2)} - \mu \right]^2 \right) \end{aligned}$$

La loi marginale de P_2 ne prend pas une forme standard, car

$$f(P_2) = \int g_0(P_1, P_2) \cdot dP_1,$$

, mais peut être obtenue de manière numérique.

Étant données les observations Z , la loi *a posteriori*

$$\begin{aligned} g(P_1, P_2|Z) &\propto h(Z|P_1, P_2)g_0(P_1, P_2) \\ &\propto h(Z|P_1, P_2) \cdot \frac{P_1^{a-1}(1-P_1)^{b-1}}{P_2(1-P_2)} \cdot \exp \left(-\frac{1}{2\sigma^2} \cdot \left[\log \frac{P_2(1-P_1)}{P_1(1-P_2)} - \mu \right]^2 \right) \end{aligned}$$

or la vraisemblance est de la forme suivante :

$$P_1^{S_1} \cdot (1-P_1)^{n_1-S_1} \cdot P_2^{S_2} \cdot (1-P_2)^{n_2-S_2}$$

donc,

$$h(Z|P_1, P_2) = P_1^{S_1} \cdot P_2^{S_2} \cdot (1-P_1)^{F_1} \cdot (1-P_2)^{F_2}$$

Par conséquent,

$$g(P_1, P_2|Z) \propto P_1^{S_1+a-1} \cdot (1-P_1)^{F_1+b-1} \cdot P_2^{S_2-1} \cdot (1-P_2)^{F_2-1} \cdot \exp \left(-\frac{1}{2\sigma^2} \cdot \left[\log \frac{P_2(1-P_1)}{P_1(1-P_2)} - \mu \right]^2 \right)$$

La distribution marginale *a posteriori* de P_2 , P_1 et θ peut être obtenue de manière numérique.

Le modèle bayésien proposé nécessite la spécification des deux lois *a priori*, c'est à dire Beta(a,b) la loi *a priori* de P_1 et $\mathcal{N}(\mu, \sigma^2)$ l'*a priori* de θ . L'opinion individuelle des experts pour ces paramètres pourrait être combinée pour dériver des distributions consensuelles en utilisant une agrégation mathématique ou comportementale. L'agrégation mathématique peut être utilisée pour combiner les distributions *a priori* de chaque expert en utilisant la méthode par moyenne des opinions ou bien la méthode de mélange des opinions, linéaire ou logarithmique, où celles-ci peuvent tenir compte des pondérations inégales des opinions des différents experts [Chapitre 3].

Comme les experts ont des expériences et connaissances semblables, sur l'efficacité du MTX par rapport au CS pour les patients atteints d'artérites à cellules géantes, nous allons utiliser dans la suite l'agrégation mathématique en utilisant la méthode par moyenne des opinions pour trouver le consensus.

Nous proposons dans la suite une approche pour déterminer comment choisir les paramètres de P_1 , P_2 et θ

4.3.3 Détermination d'une distribution *a priori* sur la base d'opinion d'experts.

Nous avons sollicité 3 experts de nationalité française et nous nous sommes réunis pour le choix des lois *a priori*.

→ Former les experts.

→ Obtenir l'opinion des experts sur P_1 , P_2 et θ

→ Obtenir l'avis des experts sur la pertinence des RCT avec

Le but de la réunion était de : les conditions relatives.

→ Combiner les avis en un consensus d'opinion

→ Vérifier la validité des distributions *a priori*

du consensus final.

Afin d'obtenir de chaque expert une distribution des paramètres P_1 , θ , nous avons établi un questionnaire de 4 questions codées sur une EVA (Echelle visuelle analogique) entre 0 et 1.

Pour déterminer les paramètres a et b de la loi Bêta de P_1 , les experts sont questionnés comme suit :

Q1 : Que pensez-vous du taux maximal des patients n'ayant présenté aucune rechute à 12 mois de suivi avec le traitement corticostéroïde ?

Q2 : Fournissez une proportion p telle que vous soyez sûr à 75% que le taux de rémission avec le traitement CS excède la valeur de cette proportion ($P(P_1 > p) = 75\%$)

La réponse à la question Q1 correspond au mode *a priori* pour P_1 , c'est-à-dire

$$\frac{a - 1}{a + b - 2}$$

et la réponse à Q2 correspond au premier quartile $\Pi_{0.25}$, tel que

$$B(\Pi_{0.25}; a, b) = 0.25$$

(B est la fonction de répartition de bêta).

Donc on obtient un système de deux équations à deux inconnues,

En résolvant ce système on obtient a et b .

Pour déterminer les paramètres μ et σ de la loi Normale de θ , les experts sont questionnés comme suit :

Q3 : Quelle est la probabilité que le taux de non-rechute à 12 mois avec le traitement MTX soit supérieur à celui obtenu avec le traitement classique CS ?

Q4 : Quelle est la probabilité que le taux de non-rechute à 12 mois avec le traitement classique CS excède celui du nouveau traitement MTX de plus de 10%.

La réponse à la question Q3 correspond à la loi de probabilité que $P_2 > P_1$ qui est ramenée à $\Phi(\mu/\sigma)$ telle que Φ corresponde à la loi normale standard $\mathcal{N}(0, 1)$.

$$P(P_2 \geq P_1) = \Phi(\mu/\sigma)$$

La question Q4 demande la probabilité *a priori* que $P_1 > P_2 + 0.1$ c'est-à-dire que le méthotrexate associé à la corticothérapie est inférieur aux corticostéroïdes seuls d'au moins la marge pré spécifiée de non-infériorité, $\delta = 0.1$.

Pour avoir une idée de l'incertitude *a priori* sur P_1 et P_2 , nous écrivons cette probabilité comme l'intégrale de la loi jointe $g_0(P_1, P_2)$.

$$P(P_2 - P_1 \leq -0.1) = \int_0^1 \int_0^{\max(P_1-0.1, 0)} g_0(P_1, P_2; a, b, \frac{\mu}{\sigma}, \sigma) dP_2 dP_1 \quad (4.2)$$

Donc,

$$\begin{aligned} P(P_2 - P_1 \geq -0.1) &= 1 - P(P_2 - P_1 \leq -0.1) \\ &= 1 - \int_0^1 \int_0^{\max(P_1-0.1, 0)} g_0(P_1, P_2; a, b, \frac{\mu}{\sigma}, \sigma) dP_2 dP_1 \end{aligned} \quad (4.3)$$

Comme cette loi jointe $g_0(P_1, P_2)$ (équation 4.1) peut être exprimée en fonction de σ et les paramètres qui sont déjà fixés par les réponses aux questions Q1 à Q3, donc le calcul de l'intégrale numérique de l'équation 4.2 permet de déterminer une valeur convenable de σ , puis de μ .

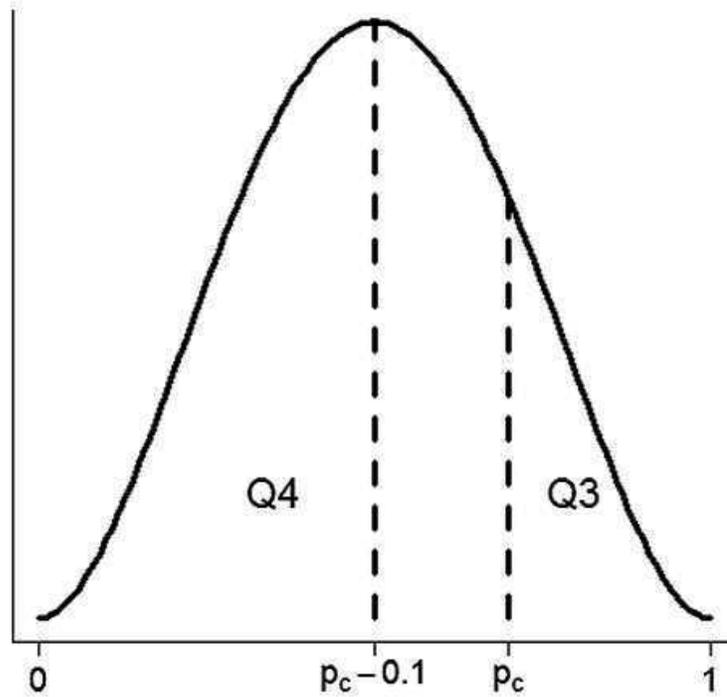


Figure 4.1 – Interprétation graphique des réponses hypothétiques aux questions Q3-Q4 [Hampson et al., 2014]

La figure 4.1 illustre l'ordre que les réponses doivent suivre ; donc la réponse à Q4 doit

être inférieure à 1- réponse à Q3.

On place les réponses aux questions Q1 à Q4 de chaque expert dans le tableau 4.3. On utilisera les "croyances" individuelles *a priori* des experts avant de les grouper.

| | Q1 | Q2 | Q3 | Q4 |
|-----------|------|------|------|------|
| Expert 1 | 0.65 | 0.55 | 0.8 | 0.1 |
| Expert 2 | 0.65 | 0.5 | 0.75 | 0.1 |
| Expert 3 | 0.7 | 0.65 | 0.55 | 0.3 |
| Consensus | 0.7 | 0.55 | 0.7 | 0.15 |

Table 4.1 – Réponses des experts aux questions Q1 à Q4 pour les études des artérites à cellules géantes

Une fois que chaque expert a rempli le questionnaire, nous avons trouvé les caractéristiques de chaque fonction de densité, y compris l'intervalle de confiance(IC) à 90%, le mode, la moyenne et la force de l'opinion *a priori*(en calculant la taille de l'échantillon efficace ESS).

En tenant en compte de la difficulté d'interpréter le log OR, nous allons interpréter la fonction de densité de θ de manière informelle en termes de distribution *a priori* de P_2 en supposant que P_1 est fixé au mode *a priori* représenté par les réponses à la question Q1.

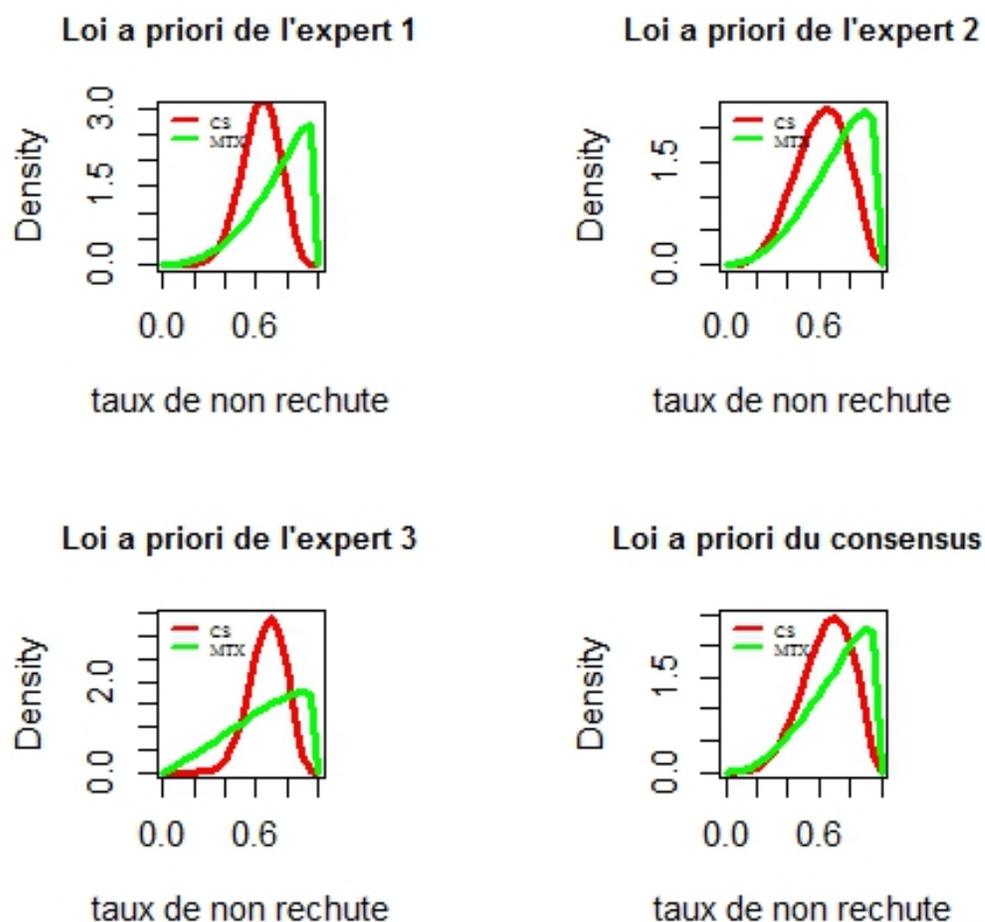


Figure 4.2 – Représentation graphique des lois *a priori* des trois experts du taux de non-rechute après 12 mois de suivi pour les patients atteints de la maladie de Horton et assignés aux CS et au MTX

4.3.4 Force de l'avis *a priori* des experts

Lors de la caractérisation de l'opinion *a priori* des experts, il leur est utile de montrer les implications de leurs réponses aux questionnaires de différentes façons.

Pour caractériser l'opinion *a priori* des experts, on calcule la taille effective de l'échantillon noté par la suite "ESS" [Morita et al., 2008].

Pour notre modèle bayésien, nous calculons les ESS, en prenant l'information préalable

pour chaque paramètre comme étant la précision de la distribution *a priori* [Neuenschwander et al., 2010]. Ce sera une approximation à moins que les lois *a priori* suivent des distributions normales.

Par conséquent, plutôt que de calculer l'ESS de la loi *a priori* Beta (α_1, β_1) avant P_1 , nous déterminons l'ESS *a priori* du log OR du taux de non-rechute dans le groupe CS, tel que $\log \text{OR} = \omega = \log \left\{ \frac{P_1}{1-P_1} \right\}$, pour lequel une approximation normale sera plus précise.

Soit

$$p_0(\omega) = \frac{\exp(a\omega)}{\left[\text{Beta}(\alpha_1, \beta_1) \{1 + \exp(\omega)\}^{\alpha_1 + \beta_1} \right]}$$

représente la loi *a priori* de ω (pour $P_1 \propto \text{Beta}(\alpha_1, \beta_1)$) et qui a une précision $1/\text{Var}_0(\omega)$. Nous définissons la quantité appropriée pour la comparaison avec $1/\text{Var}_0(\omega)$ comme l'information *a priori* attendue pour ω qui serait générée en effectuant un essai à un seul bras traitant n patients avec CS.

Dans cet esprit, soit $V_\omega(n)$ représente l'information attendue de Fisher pour ω qui serait accumulée lors d'un essai à un seul bras lorsque les distributions des patients sont distribuées selon la loi Bernouilli, $Y_{CSi} \propto \text{Bernouilli}(P_1)$, avec $i = 1, \dots, n$.

Alors

$$V_\omega(n) = \frac{n \exp(\omega)}{\{1 + \exp(\omega)\}^2},$$

et l'*a priori* $E_0[V_\omega(n)]$ de $V_\omega(n)$, est calculée en intégrant l'information de la loi *a priori* de ω .

Nous définissons l'ESS de $p_0(\omega)$ comme la taille n^* , pour laquelle $E_0[V_\omega(n^*)]$ égale la précision $1/\text{Var}_0(\omega)$, tel que n^* doit satisfaire :

$$n^* \int_{-\infty}^{\infty} \frac{1}{\text{Beta}(\alpha_1, \beta_1)} \frac{\exp\{(\alpha_1 + 1)\omega\}}{\{1 + \exp(\omega)\}^{\alpha_1 + \beta_1 + 2}} d\omega = \frac{1}{\text{var}_0(\omega)}$$

D'autre part, il nous faut calculer l'ESS de $k_0(\theta)$, comme θ mesure l'avantage de E sur C, on définit l'ESS de $k_0(\theta)$ comme la taille n_θ^* de l'échantillon pour laquelle

$$\frac{1}{\text{Var}_0(\theta)} = E_0[V_\theta(n_\theta^*)];$$

tel que n_θ^* doit satisfaire :

$$\int_0^1 \int_0^1 \text{Var}_\theta(n_\theta^*) g_0(P_2, P_1) dP_2 P_1 = \frac{1}{\sigma^2}.$$

4.3.5 Implémentation numérique sur le logiciel R

Les caractéristiques des lois *a priori* élicitées de P_1 , P_2 et θ ont été générées avec le logiciel R en utilisant le package "Shiny".

En se basant sur les réponses présentées dans le tableau 4.3 on présente ci-dessous les fonctions de densité présentant l'avis des experts sur le taux de non-rechute à 12 mois de suivi du CS et du MTX .

- L'*a priori* de P_1 est une loi Beta(5.64;2.99), de moyenne égale à 0.65, mode = 0.7, un écart-type = 0.15 et IC à 90% = [0.38;0.88], la taille ESS de l'*a priori* du $\log \left\{ \frac{P_1}{1-P_1} \right\}$ est de 8 patients dans le groupe CS, donc le nombre de patients nécessaire pour obtenir la même quantité d'informations statistique représentée par la loi *a priori* de P_1 est de 8 patients.

- L'*a priori* de θ est une loi normale $\mathcal{N}(0.5;0.92)$. de moyenne égale à 0.5, mode = 0.5, un écart-type = 0.96 et IC à 90% = [-1.08; 2.08] La taille ESS de la loi *a priori* de θ est égale à 11 patients pour chaque groupe de traitement, donc le nombre de patients nécessaire pour obtenir la même quantité d'information statistique représentée par la loi *a priori* de θ est de 11 patients pour chacun des groupes MTX et CS.

À partir des paramètres trouvés P_1 et θ et à partir du modèle proposé, l'*a priori* de P_2 est une loi Beta(3.12;1.18), de moyenne égale à 0.72, mode = 0.92, un écart-type = 0.2 et IC à 90% = [0.31;0.96], la taille ESS de l'*a priori* du $\log \left\{ \frac{P_2}{1-P_2} \right\}$ est de 4 patients dans le groupe MTX.

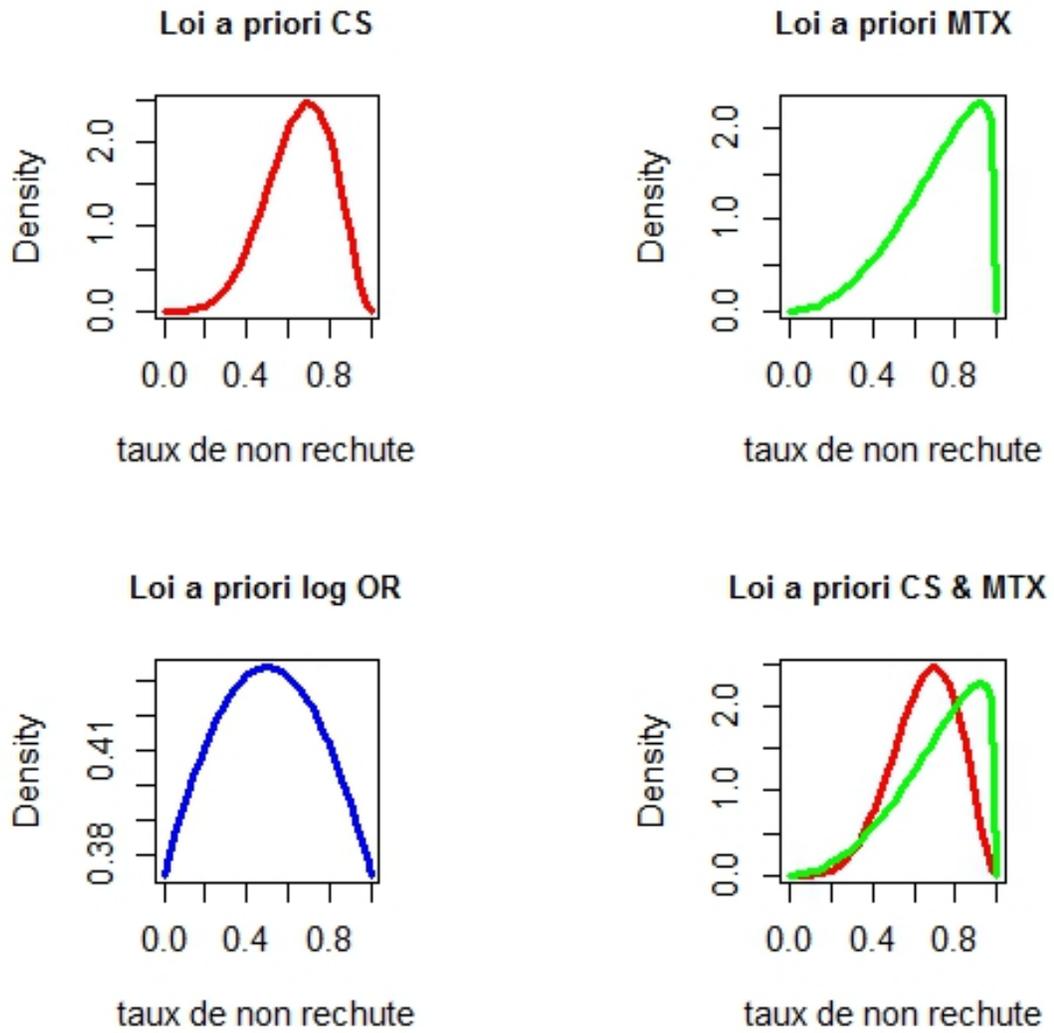


Figure 4.3 – Représentation graphique de la loi a priori du log OR et des lois a priori du taux de non-rechute après 12 mois de suivi pour les patients atteints de la maladie de Horton et assignés aux CS ou au MTX

4.3.6 Loi *a posteriori* pour un échantillon fixé

Une fois que la distribution *a priori* a été établie, nous considérons pour chaque taille d'échantillon, les valeurs possibles (S_1, F_1, S_2, F_2) , avec S correspondant au nombre de succès et F au nombre d'échecs. Les distributions *a posteriori* correspondantes peuvent alors être affichées. Par exemple, si on présente un échantillon de taille $n=40$, on a alors 20 patients dans chaque groupe de traitement, tel qu'on considère que P_1 est fixé au mode, donc

$$P_1 = \frac{S_1}{n_1} = 0.7 \Rightarrow S_1 = 0.7 * 20 = 14.$$

En fixant le taux de non-rechute du bras CS à 14 patients, nous considérons les cas pour lesquelles le taux de non-rechute du MTX est équivalent, juste inférieur et très inférieur de celui du CS, d'où, $S_2 = 14$, $S_2 = 12$ et $S_2 = 10$ respectivement.

Pour le cas de $S_2 = 14$ on obtient les *a posteriori* suivants :

- La distribution *a posteriori* de P_1 a comme moyenne= 0.67, mode= 0.69, écart-type=0.08, IC à 90%= [0.54; 0.8]
- La distribution *a posteriori* de P_2 a comme moyenne= 0.72, mode= 0.74, écart-type=0.08, IC à 90%= [0.56; 0.85]
- La distribution *a posteriori* de θ a comme moyenne= 0.22, mode= 0.22, écart-type=0.49, IC à 90%= [-0.67; 1.01]

Donc dans cet échantillon ($(S_1 = 14, F_1 = 6, S_2 = 14, F_2 = 6)$) nous obtenons :

$$P(P_2 - P_1 \geq -0.1|z) = 0.9$$

Pour le cas de $S_2 = 12$ on obtient les *a posteriori* suivants :

- La distribution *a posteriori* de P_1 a comme moyenne= 0.65, mode= 0.66, écart-type=0.08, IC à 90%= [0.52; 0.79]
- La distribution *a posteriori* de P_2 a comme moyenne= 0.64, mode= 0.65, écart-type=0.09, IC à 90%= [0.47; 0.79]
- La distribution *a posteriori* de θ a comme moyenne=-0.12, mode= -0.13, écart-type=0.52, IC à 90%= [-0.98; 0.74]

Donc dans cet échantillon $((S_1 = 14, F_1 = 6, S_2 = 12, F_2 = 8))$ nous obtenons :

$$P(P_2 - P_1 \geq -0.1|z) = 0.73$$

Pour le cas de $S_2 = 10$ on obtient les *a posteriori* suivants :

- La distribution *a posteriori* de P_1 a comme moyenne= 0.65, mode= 0.66, écart-type=0.08, IC à 90%= [0.51; 0.78]
- La distribution *a posteriori* de P_2 a comme moyenne= 0.55, mode= 0.56, écart-type=0.1, IC à 90%= [0.38; 0.71]
- La distribution *a posteriori* de θ a comme moyenne=-0.43, mode= -0.44, écart-type=0.51, IC à 90%= [-1.28; 0.41]

Donc dans cet échantillon $((S_1 = 14, F_1 = 6, S_2 = 10, F_2 = 10))$ nous obtenons :

$$P(P_2 - P_1 \geq -0.1|z) = 0.5$$

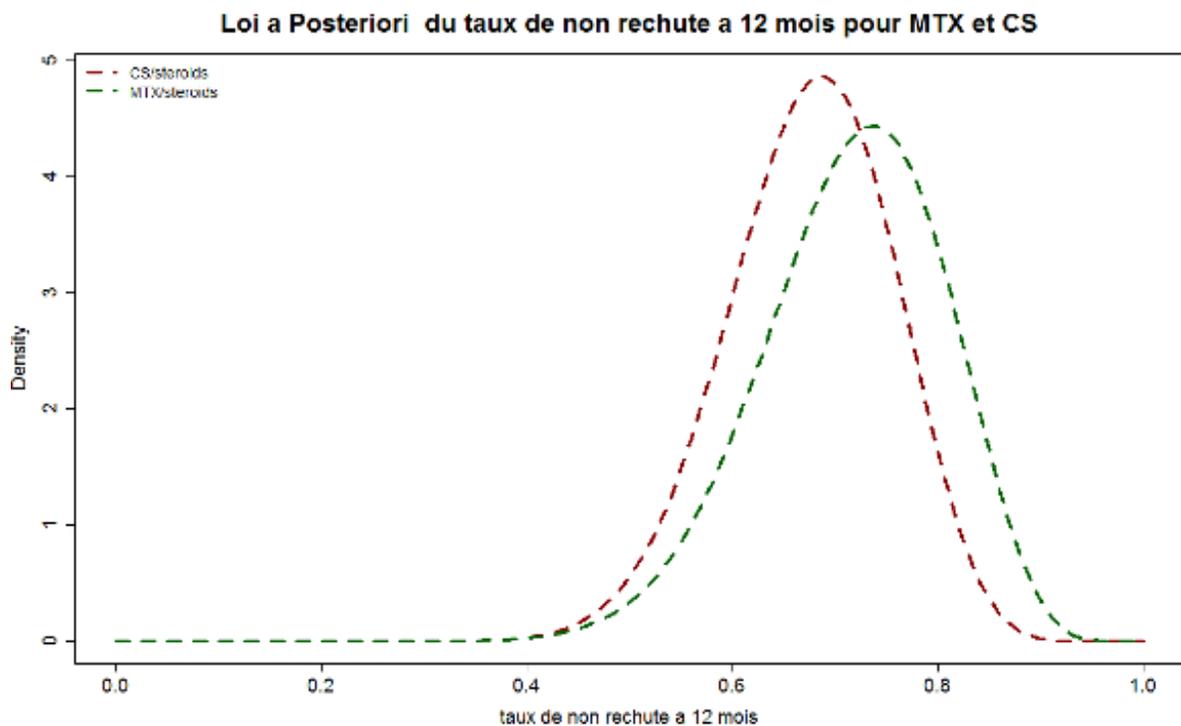


Figure 4.4 – Représentation graphique de la loi a priori du taux de non-rechute après 12 mois de suivi pour l'échantillon $S_2 = 14$

4.3.6.1 Interprétation

Nous constatons que les probabilités d'avoir une non-infériorité avec un seuil de 10% diminuent au fur et à mesure que le nombre de succès dans le groupe méthotrexate associé aux corticoïdes diminue, ce qui est logique.

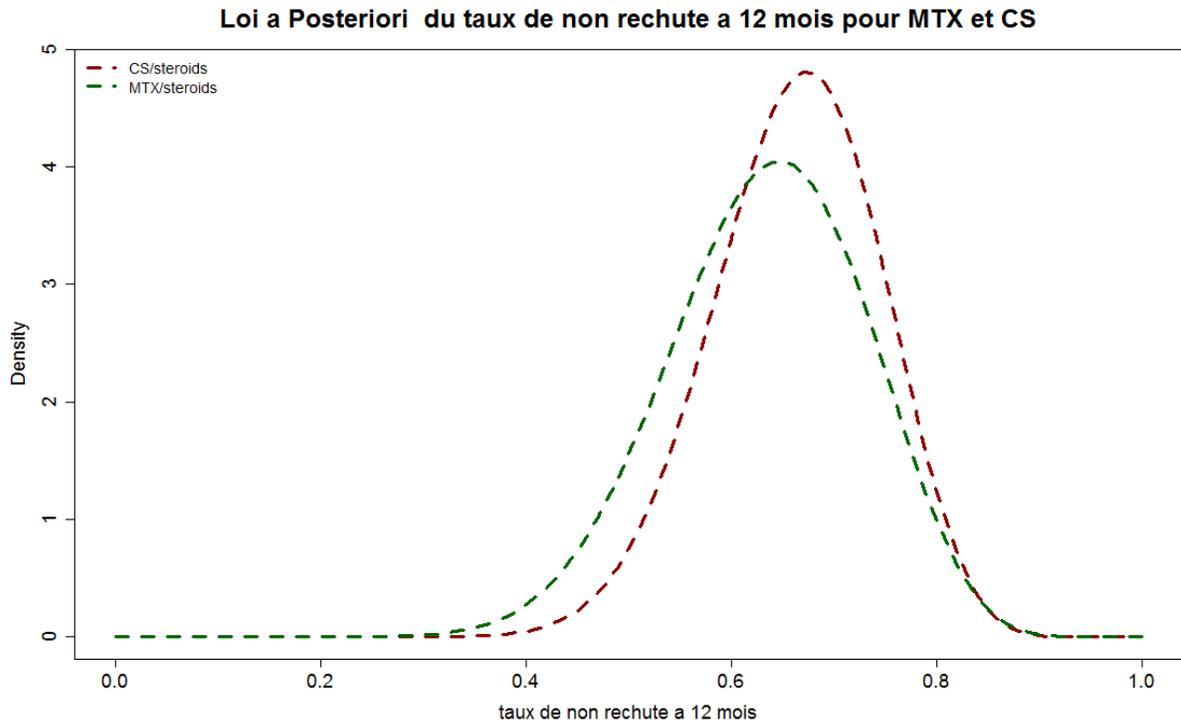


Figure 4.5 – Représentation graphique de la loi a priori du taux de non-rechute après 12 mois de suivi pour l'échantillon $S_2 = 12$

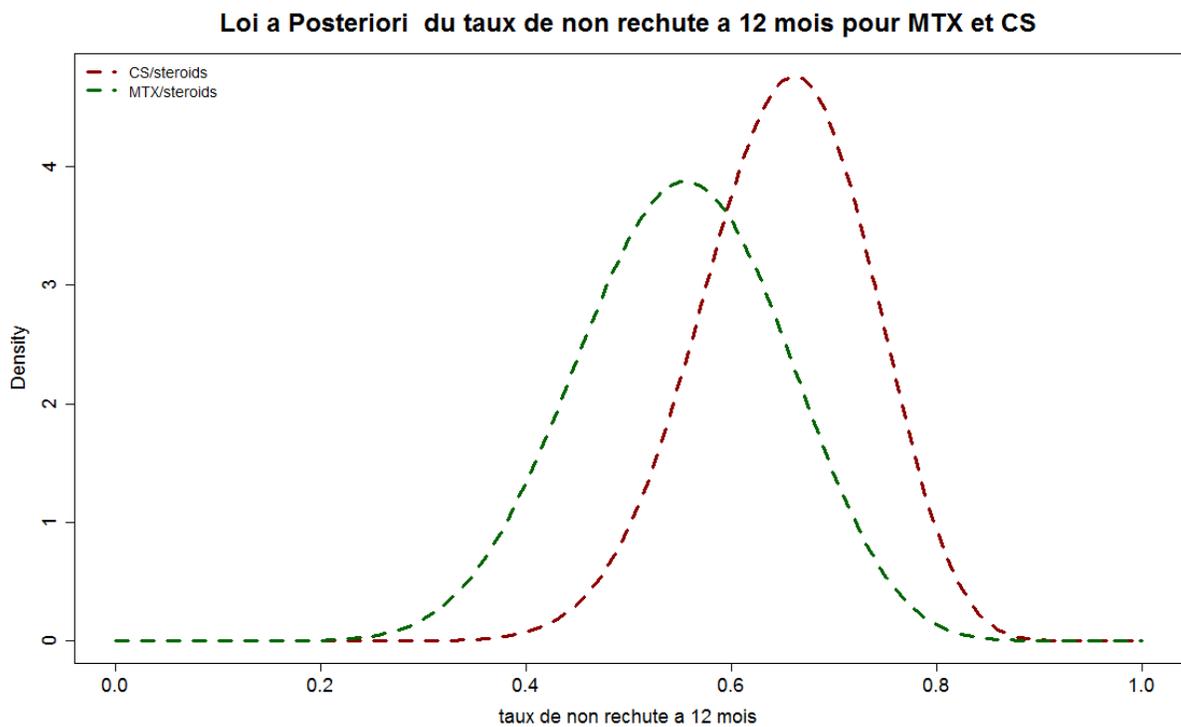


Figure 4.6 – Représentation graphique de la loi a priori du taux de non-rechute après 12 mois de suivi pour l'échantillon $S_2 = 10$

4.3.7 Probabilité de non-infériorité en incorporant des données d'études historiques

Nous allons maintenant calculer la probabilité de non-infériorité, tel que le taux des patients ne représentant aucune rechute dans le groupe MTX associé aux corticoïdes soit supérieur à celui du groupe CS moins un seuil de 10%, en incorporant les données de trois études randomisées ($\Leftrightarrow P(P_2 - P_1 \geq -0.1|Z)$, avec $Z =$ données provenant des études). Nous présentons alors dans le tableau 4.2 trois études randomisées sur l'efficacité du MTX associé aux corticoïdes vs CS seuls pour la maladie de Horton

| Étude | Essai | Populations | date de suivi | Patient avec aucune rechute | Durée et dose de dose cumulée du CS |
|------------------------|---|-----------------|---------------|-----------------------------|---------------------------------------|
| [Spiera et al., 2001] | A multicenter, randomized, double-blind, placebo-controlled trial | 12 MTX vs 9 CS | NA | 50 % vs 67 % P= NA | 68 sem vs 60 sem ; 6.5 g vs 5.9 g |
| [Jover et al., 2001] | Single center double blinded | 21 MTX vs 21 CS | 24 mois | 55% vs 16% P=0.004 | 29 sem vs 94 sem ; 4.3 g vs 5.4 g |
| [Hoffman et al., 2002] | A multicenter, randomized, double-blind, placebo-controlled trial | 51 MTX vs 47 CS | 12 mois | 39% vs 34% , P= 0.31 | 5.4 mois vs 5.6 mois ; 5.4 g vs 5.3 g |

Table 4.2 – Trois essais randomisés de GCA

Pour l'essai présenté par [Spiera et al., 2001], nous avons

$$P_1 = \frac{S_1}{n_1} \Rightarrow S_1 = 0.67 * 9 = 6$$

et

$$P_2 = \frac{S_2}{n_2} \Rightarrow S_2 = 0.5 * 12 = 6$$

Donc dans ce cas à partir de l'équation suivante

$$\begin{aligned} P(P_2 - P_1 \geq -0.1) &= g(P_1, P_2 | Z_{Spiera}) \\ &\propto P_1^{S_1+a-1} \cdot (1 - P_1)^{F_1+b-1} \cdot P_2^{S_2-1} \cdot (1 - P_2)^{F_2-1} \cdot \exp\left(-\frac{1}{2\sigma^2} \cdot \left[\log \frac{P_2(1-P_1)}{P_1(1-P_2)} - \mu\right]^2\right) \\ &= 0.64 \end{aligned}$$

Pour l'essai présenté par [Jover et al., 2001] nous avons

$$P_1 = \frac{S_1}{n_1} \Rightarrow S_1 = 0.16 * 21 = 3$$

et

$$P_2 = \frac{S_2}{n_2} \Rightarrow S_2 = 0.55 * 21 = 12$$

Donc dans ce cas à partir de l'équation suivante

$$\begin{aligned} P(P_2 - P_1 \geq -0.1) &= g(P_1, P_2 | Z_{Jover}) \\ &= 0.99 \end{aligned}$$

Pour l'essai présenté par [Hoffman et al., 2002] nous avons

$$P_1 = \frac{S_1}{n_1} \Rightarrow S_1 = 0.34 * 47 = 16$$

et

$$P_2 = \frac{S_2}{n_2} \Rightarrow S_2 = 0.39 * 51 = 20$$

Donc dans ce cas à partir de l'équation suivante

$$\begin{aligned} P(P_2 - P_1 \geq -0.1) &= g(P_1, P_2 | Z_{Hoffman}) \\ &= 0.92 \end{aligned}$$

4.3.7.1 Interprétation

Nous remarquons donc que pour les trois études, la probabilité de non-infériorité au seuil de 10% est supérieure à 90% dans le cas des études [Jover et al., 2001] et [Hoffman et al., 2002] qui représentent une efficacité du méthotrexate associé aux corticoïdes par rapport aux corticoïdes seuls. D'autre part, même dans le cas de l'essai de [Spiera et al., 2001] où ils représentent une efficacité (non significative) des CS par rapport au MTX associé aux corticoïdes, la probabilité d'acceptation de la non-infériorité

reste en faveur du traitement MTX (64%).

4.4 Deuxième méthode : Méthodes de différences de deux lois bêta

4.4.1 Principe de la méthode

Soit

$$B_\theta(a, b) = \frac{\theta^{a-1}(1-\theta)^{b-1}}{B(a, b)} \quad \text{avec } \theta \in [0, 1]$$

avec a et b les paramètres de la loi bêta et la fonction bêta B(a,b) présentée par :

$$B(a, b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$$

où $\Gamma(a)$ est la valeur de la loi Γ au point a.

Dans le cas où on a deux variables aléatoires, par exemple, P_1 et P_2 qui sont indépendantes

.

On peut chercher à estimer :

$$P(P_2 - P_1 \geq \delta),$$

c'est-à-dire la probabilité que le taux de succès dans le bras traitement excède celle du bras témoin d'un seuil δ , On pose alors,

$$g(a, b, c, d, \delta) = P(P_2 - P_1 \geq \delta), \text{ avec } \delta = -10\%$$

avec, $P_1 \propto \text{Beta}(a,b)$ et $P_2 \propto \text{Beta}(c,d)$, avec les paramètres a, b, c, d ≥ 0 :

$$\begin{aligned} g(a, b, c, d, \delta) &= P(P_2 - P_1 \geq \delta) \\ &= \int_0^1 f_2(P_2) I_{P_2-\delta} dP_1 dP_2 \\ &= \int_0^1 \left(\int_0^{P_2-\delta} \frac{P_1^{a-1}(1-P_1)^{b-1}}{B(a, b)} dP_1 \right) \frac{P_2^{c-1}(1-P_2)^{d-1}}{B(c, d)} dP_2 \\ &= \frac{1}{B(a, b) * B(c, d)} \int_0^1 \left(\int_0^{P_2+0.1} P_1^{a-1}(1-P_1)^{b-1} dP_1 \right) P_2^{c-1}(1-P_2)^{d-1} dP_2 \end{aligned} \quad (4.4)$$

4.4.2 Choix des *a priori* en se basant sur P1 et P2

Nous présentons dans cette partie une autre méthode pour réaliser le test de non-infériorité du MTX par rapport au CS. Tout d'abord on présente les deux questions

suivantes aux trois experts afin de calculer les paramètres de la loi Beta de P_2 .

Q5 : Que pensez-vous du taux maximal de patients n'ayant aucune rechute à 12 mois de suivi avec le traitement MTX ?

Q6 : Fournissez une proportion p telle que vous soyez sûr à 75% que le taux de rémission avec le traitement MTX excède la valeur de cette proportion ($P(P_2 > p) = 75\%$).

La réponse à la question Q5 correspond au mode *a priori* pour p_2 , c'est-à-dire

$$\frac{c - 1}{c + d - 2}$$

et la réponse à Q6 correspond au premier quartile $\Pi_{0.25}$, tel que

$$B(\Pi_{0.25}; c, d) = 0.25$$

(B est la fonction de répartition de bêta).

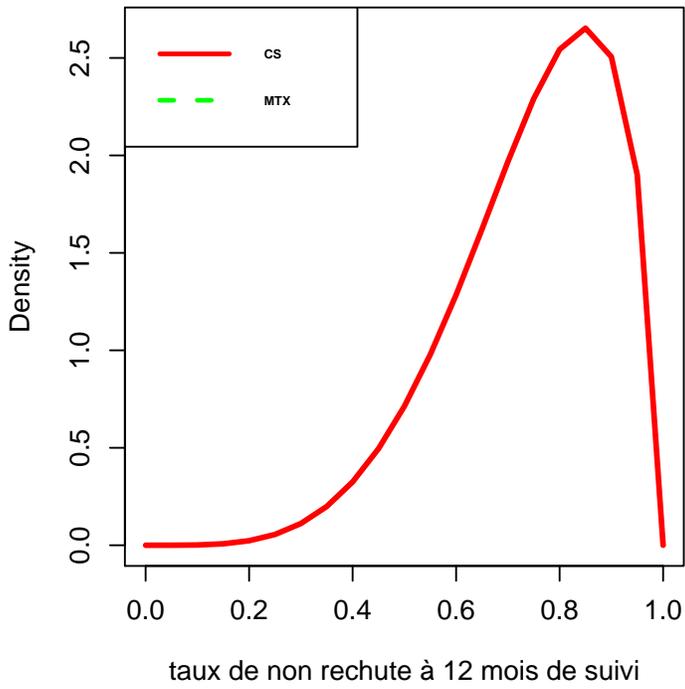
On place les réponses aux questions Q1, Q2, Q5 et Q6 de chaque expert dans le tableau 4.3. On utilisera les "croyances" individuelles *a priori* des experts avant de les grouper.

| | Q1 | Q2 | Q5 | Q6 |
|-----------|------|------|------|------|
| Expert 1 | 0.85 | 0.65 | 0.9 | 0.7 |
| Expert 2 | 0.65 | 0.5 | 0.8 | 0.55 |
| Expert 3 | 0.7 | 0.5 | 0.8 | 0.55 |
| Consensus | 0.7 | 0.55 | 0.85 | 0.6 |

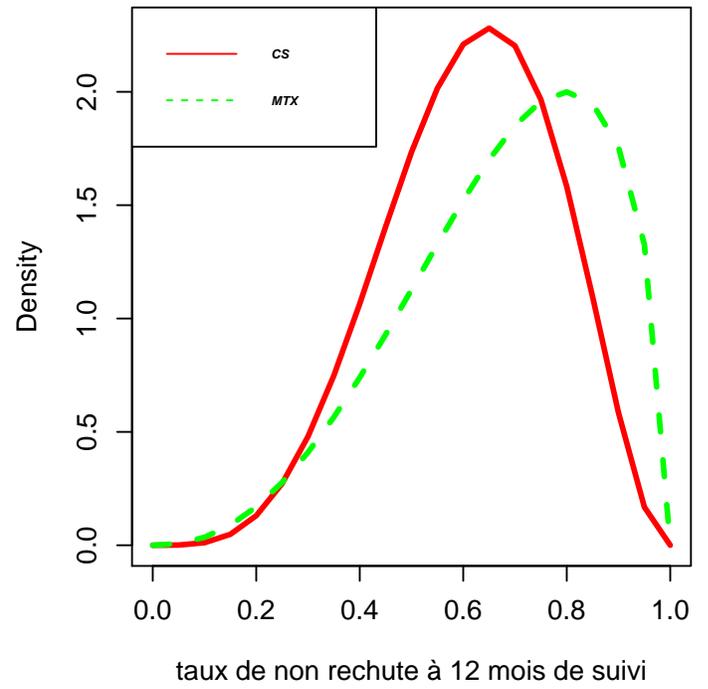
Table 4.3 – Réponses des experts aux questions Q1, Q2, Q3 et Q4 pour les études relatives à l'artérite à cellules géantes

En utilisant le logiciel R et en se basant sur les réponses présentées dans le tableau 4.3 on présente ci-dessous les fonctions de densité présentant l'avis de chaque expert sur le taux de non-rechute à 12 mois de suivi du CS & MTX

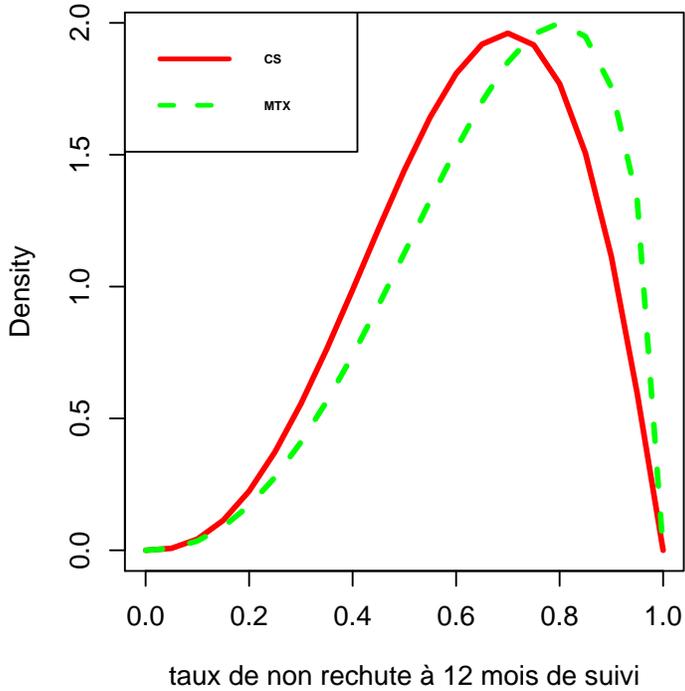
Loi a priori de l'expert 1



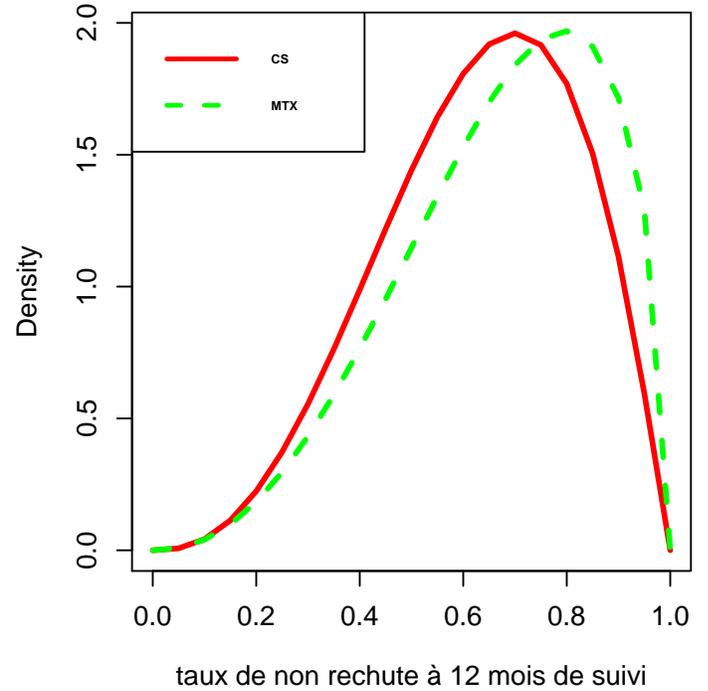
Loi a priori de l'expert 2



Loi a priori de l'expert 3



Loi a priori du consensus



Les réponses aux questions Q1 à Q4 permettent de définir les paramètres des lois de p_C et de θ :

$$P_1 \hookrightarrow \text{Beta}(5.64, 2.99)$$

$$P_2 \hookrightarrow \text{Beta}(3.77, 1.49)$$

4.4.3 Calcul de la loi *a posteriori*

Nous avons présenté dans le "Chapitre 2 exemple 3.1" le modèle bayésien pour le calcul de la distribution *a posteriori* bêta.

Comme p_1 et p_2 suivent des lois $\text{Beta}(\alpha_1, \beta_1)$ et $\text{Beta}(\alpha_2, \beta_2)$ respectivement et dans notre modèle la vraisemblance correspond à la probabilité d'observer x_1, x_2 succès (succès=patients sans rechute) conditionnellement à p_1 et p_2 respectivement. Donc les vraisemblances suivent des lois binomiales, $B(n_1, p_1)$ et $B(n_2, p_2)$ respectivement.

Donc les lois *a posteriori* sont aussi des lois Bêta :

$$\pi(p_1|x_1) \hookrightarrow \text{Beta}(x_1 + \alpha_1, n_1 - x_1 + \beta_1)$$

et

$$\pi(p_2|x_2) \hookrightarrow \text{Beta}(x_2 + \alpha_2, n_2 - x_2 + \beta_2) \text{ respectivement}$$

avec n_1, n_2 taille des échantillons de p_1 et p_2 respectivement.

Le but de notre modèle est d'étudier la probabilité que le taux de succès dans le bras MTX excède celui dans le bras CS d'au moins $\delta = -10\%$ (valeur définie par les cliniciens).

4.4.4 Incorporation des données d'un essai voisin et résultats

On présente un essai randomisé publié dans [Jover et al., 2001] de $n=42$ patients atteints d'artérite à cellules géantes. Les patients sont groupés en deux groupes, le premier groupe recevant le methotrexate et le deuxième groupe recevant la corticothérapie seule. Les patients sont âgés de plus de 50 ans avec une majorité de femmes (66.7% dans le groupe MTX et 71.4% dans le groupe CS). Pour 24 mois de suivi, 2 patients n'ont pas continué le suivi dans le groupe CS et 1 seul patient n'a pas continué les 24 mois de suivi dans le groupe MTX.

À la fin du suivi, 3 patients, parmi les 19 patients qui ont complété le suivi, n'ont eu aucune rechute dans le groupe CS (15.8%), et 11 patients, parmi les 20 qui ont complété le suivi, n'ont eu aucune rechute dans le groupe MTX (55%).

Donc $P'_1 = \pi(p_1|x_1) \hookrightarrow \text{Beta}(5.64 + 3, 2.99 + 19 - 3) \Rightarrow \pi(p_1|x_1) \hookrightarrow \text{Beta}(8.64, 18.99)$,
et $P'_2 = \pi(p_2|x_2) \hookrightarrow \text{Beta}(3.77 + 11, 1.49 + 20 - 11) \Rightarrow \pi(p_1|x_1) \hookrightarrow \text{Beta}(14.77, 10.49)$,

Alors à partir des lois *a posteriori* calculées ci-dessus : $\alpha'_1 = 8.64$, $\alpha'_2 = 14.77$, $\beta'_1 = 18.99$, $\beta'_2 = 10.49$, et en appliquant la fonction $g(\alpha_1, \beta_1, \alpha_2, \beta_2, \delta)$, on peut calculer maintenant ,

$$\mathcal{P}(P'_2 - P'_1 \geq \delta|x) = g(\alpha'_1, \beta'_1, \alpha'_2, \beta'_2, \delta),$$

avec $P'_1 = P_1|x_1$ et $P'_2 = P_2|x_2$

Donc, comme $P_1|x_{1Jover}$ et $P_2|x_{2Jover}$ suivent des lois bêta, nous pouvons appliquer l'équation 4.4 dans le but de calculer $P(P_2 - P_1 \geq \delta|x)$ alors,

$$\begin{aligned} P(P_2 - P_1 \geq \delta|x_{Jover}) &= g(a', b', c', d', \delta) \\ &= \frac{1}{B(a', b') * B(c', d')} \int_0^1 \left(\int_0^{P_2 - \delta} P_1^{a'-1} (1 - P_1)^{b'-1} dP_1 \right) P_2^{c'-1} (1 - P_2)^{d'-1} dP_2 \\ &= \frac{1}{\text{Beta}(14.77, 10.49) * \text{Beta}(8.64, 18.99)} \\ &\int_0^1 \left(\int_0^{P_2 + 0.1} P_1^{8.64-1} (1 - P_1)^{18.99-1} dP_1 \right) P_2^{14.77-1} (1 - P_2)^{10.49-1} dP_2 \\ &= K * \left(\int_0^{P_2 + 0.1} P_1^{7.64} (1 - P_1)^{17.99} dP_1 \right) P_2^{13.77} (1 - P_2)^{9.49} dP_2 \\ &= K * (1.33e - 15) \\ &= 0.99 \end{aligned} \tag{4.5}$$

Avec les données de l'essai présenté par [Hoffman et al., 2002] nous obtenons,

$$P'_1 = \pi(p_1|x_{1_{Hoffman}}) \hookrightarrow \text{Beta}(21.64, 33.99),$$

et,

$$P'_2 = \pi(p_2|x_{2_{Hoffman}}) \hookrightarrow \text{Beta}(14.77, 10.49)$$

$$\begin{aligned}
 P(P_2 - P_1 \geq \delta|x_{Hoffman}) &= g(a', b', c', d', \delta) \\
 &= \frac{1}{B(a', b') * B(c', d')} \int_0^1 \left(\int_0^{P_2+\delta} P_1^{a'-1} (1 - P_1)^{b'-1} dP_1 \right) P_2^{c'-1} (1 - P_2)^{d'-1} dP_2 \\
 &= \frac{1}{\text{Beta}(14.77, 10.49) * \text{Beta}(21.64, 33.99)} \\
 &\int_0^1 \left(\int_0^{P_2+\delta} P_1^{21.64-1} (1 - P_1)^{33.99-1} dP_1 \right) P_2^{14.77-1} (1 - P_2)^{10.49-1} dP_2 \\
 &= K * \left(\int_0^{P_2+\delta} P_1^{20.64} (1 - P_1)^{32.99} dP_1 \right) P_2^{13.77} (1 - P_2)^{9.49} dP_2 \\
 &= K * (7.11e - 34) \\
 &= 0.92
 \end{aligned} \tag{4.6}$$

Avec les données de l'essai présenté par [Spiera et al., 2001], nous obtenons,

$$P'_1 = \pi(p_1|x_{1_{Spiera}}) \hookrightarrow \text{Beta}(14.64, 5.99),$$

et,

$$P'_2 = \pi(p_2|x_{2_{Spiera}}) \hookrightarrow \text{Beta}(9.77, 7.49)$$

$$\begin{aligned}
P(P_2 - P_1 \geq \delta | x_{Spiera}) &= g(a', b', c', d', \delta) \\
&= \frac{1}{B(a', b') * B(c', d')} \int_0^1 \left(\int_0^{P_2 + \delta} P_1^{a' - 1} (1 - P_1)^{b' - 1} dP_1 \right) P_2^{c' - 1} (1 - P_2)^{d' - 1} dP_2 \\
&= \frac{1}{Beta(9.77, 7.49) * Beta(14.64, 5.99)} \\
&\int_0^1 \left(\int_0^{P_2 + \delta} P_1^{14.64 - 1} (1 - P_1)^{5.99 - 1} dP_1 \right) P_2^{9.77 - 1} (1 - P_2)^{7.49 - 1} dP_2 \\
&= K * \left(\int_0^{P_2 + \delta} P_1^{13.64} (1 - P_1)^{4.99} dP_1 \right) P_2^{8.77} (1 - P_2)^{6.49} dP_2 \\
&= K * (1.744e - 11) \\
&= 0.38
\end{aligned} \tag{4.7}$$

4.4.5 Interprétation

Dans les deux cas [Jover et al., 2001] [Hoffman et al., 2002] où la puissance est suffisante par rapport à l'écart obtenu, les probabilités *a posteriori* sont élevées et concordent avec les valeurs obtenues avec la méthode précédente.

Par contre, pour [Spiera et al., 2001] elle est égale à 38%, ce qui est différent des 64% trouvés par la méthode précédente, mais est cependant conforme à l'inversion de l'efficacité retrouvée dans cette étude.

4.5 Troisième méthode : Méthodes ϵ -contamination

4.5.1 Principe de la méthode

Dans l'analyse bayésienne, une méthode importante pour modéliser l'incertitude sur la loi *a priori* est la méthode ϵ contamination.

On suppose que la loi *a priori* appartient à une famille de lois qui s'écrit sous la forme :

$$\pi(\theta) = (1 - \epsilon)\pi_0(\theta) + \epsilon\pi_1(\theta) \quad [\text{Berger, 1985a}]$$

Dans notre problématique, $\Pi_0(\theta)$ correspond à la loi de θ issue de l'étude la plus proche de la question posée et $\Pi_1(\theta)$ correspond à une autre étude plus "éloignée". En plus, ϵ reflète combien on pense que π doit être proche de π_0 , avec $0 < \epsilon < 1$.

Ces familles de contaminations incluent :

- Toutes les contaminations possibles,
- Toutes les contaminations symétriques et unimodales.
- Toutes les contaminations telles que π_1 est unimodale.

Avant de commencer à présenter les caractéristiques de cette méthode, on suppose X la variable aléatoire observable qui sera supposée avoir une densité $f(x|\theta)$, où θ est le paramètre inconnu appartenant à l'ensemble des paramètres Θ . Une distribution *a priori* sur Θ est notée par $\Pi(\theta)$, d'où la distribution marginale de X est donnée par :

$$m_\pi(x) = E^\pi f(x|\theta) = \int f(x, \theta)d\theta = \int f(x|\theta)\pi(\theta)d\theta$$

or,

$$\begin{aligned} \int f(x|\theta)\Pi(\theta)d\theta &= \int f(x|\theta)[(1 - \epsilon)\pi_0(\theta) + \epsilon\pi_1(\theta)]d\theta \\ &= (1 - \epsilon)\left(\int f(x|\theta)\pi_0(\theta)d\theta\right) + \epsilon\left(\int f(x|\theta)\pi_1(\theta)d\theta\right) \end{aligned}$$

d'où

$$m_\pi(x) = (1 - \epsilon)m_{\pi_0}(x) + \epsilon m_{\pi_1}(x) \tag{4.8}$$

La distribution *a posteriori* de θ étant donné x , est définie par $\pi(\cdot|x)$ et est définie par :

$$\begin{aligned}\pi(\theta|x) &= \frac{f(x, \theta)}{m_\pi(x)} \\ &= \frac{f(x|\theta) \cdot \pi(\theta)}{m_\pi(x)}\end{aligned}\tag{4.9}$$

donc d'après l'équation 4.9 on obtient :

$$\begin{cases} \pi_0(\theta|x) = \frac{f(x|\theta) \cdot \pi_0(\theta)}{m_{\pi_0}(x)} \\ \pi_1(\theta|x) = \frac{f(x|\theta) \cdot \pi_1(\theta)}{m_{\pi_1}(x)} \end{cases}$$

Proposition 4.5.1. *On suppose que $\pi = (1 - \epsilon)\pi_0 + \epsilon\pi_1$, et que les distributions *a posteriori* π_0 et π_1 existent, et que $m_\pi(x) > 0$, alors :*

$$\pi(\theta|x) = \lambda(x) \cdot \pi_0(\theta|x) + [1 - \lambda(x)] \pi_1(\theta|x)$$

$$\text{avec : } \lambda(x) = \frac{(1-\epsilon)m_{\pi_0}(x)}{m_\pi(x)}$$

Preuve:

$$\begin{aligned}\pi(\theta|x) &= \frac{f(x|\theta) \cdot \pi(\theta)}{m_\pi(x)} \\ &= \frac{f(x|\theta)}{m_\pi(x)} [(1 - \epsilon)\pi_0(\theta) + \epsilon\pi_1(\theta)] \\ &= \frac{1}{m_\pi(x)} [(1 - \epsilon)\pi_0(\theta)f(x|\theta) + \epsilon\pi_1(\theta)f(x|\theta)] \\ &= \frac{1}{m_\pi(x)} [(1 - \epsilon)m_{\pi_0}(x) \cdot \pi_0(\theta|x) + \epsilon m_{\pi_1}(x) \cdot \pi_1(\theta|x)] \\ &= \pi_0(\theta|x) [(1 - \epsilon) \frac{m_{\pi_0}(x)}{m_\pi(x)}] + \pi_1(\theta|x) [\epsilon \frac{m_{\pi_1}(x)}{m_\pi(x)}]\end{aligned}$$

or d'après l'équation 4.8 nous obtenons :

$$\epsilon m_{\pi_1}(x) = m_\pi(x) - (1 - \epsilon)m_{\pi_0}(x)$$

donc :

$$\frac{\epsilon m_{\pi_1}(x)}{m_\pi(x)} = 1 - (1 - \epsilon) \frac{m_{\pi_0}(x)}{m_\pi(x)}$$

Alors : $\pi(\theta|x) = \lambda(x) \cdot \pi_0(\theta|x) + [1 - \lambda(x)] \pi_1(\theta|x)$

$$\text{Pour } \lambda(x) = \frac{(1-\epsilon)m_{\pi_0}(x)}{m_\pi(x)}$$

Nous obtenons la distribution *a posteriori* suivante :

$$\pi(\theta|x) = \lambda(x) \cdot \pi_0(\theta|x) + [1 - \lambda(x)] \pi_1(\theta|x)$$

□

4.5.2 Risque bayésien

Dans le problème de décision, dans le but d'avoir une comparaison effective des lois choisies, on cherche à minimiser le coût ou la perte moyenne $\rho(\pi(\theta|x), a)$ présentée comme suit :

$$\begin{aligned}
 \rho(\pi(\theta|x), a) &= \mathbb{E}^{\pi(\theta|x)}.L(\theta, a) \\
 &= \int L(\theta, a).\pi(\theta|x)d\theta \\
 &= \int L(\theta, a) [\lambda(x).\pi_0(\theta|x) + [1 - \lambda(x)]\pi_1(\theta|x)] d\theta \\
 &= \int L(\theta, a)\lambda(x).\pi_0(\theta|x)d\theta + \int L(\theta, a)[1 - \lambda(x)]\pi_1(\theta|x)d\theta \\
 &= \lambda(x)\rho(\pi_0(\theta|x), a) + (1 - \lambda(x))\rho(\pi_1(\theta|x), a)
 \end{aligned} \tag{4.10}$$

4.5.2.1 Risque bayésien avec fonction de coût quadratique

On rappelle qu'avec une fonction de perte quadratique, l'estimateur bayésien est la moyenne de la loi *a posteriori* $r(\pi) = \int \theta\pi(\theta|x)d\theta$.

Mais

$$\pi(\theta|x) = \lambda(x).\pi_0(\theta|x) + [1 - \lambda(x)]\pi_1(\theta|x) \text{ (Proposition 4.5.1)}$$

où $\pi_0(\theta|x)$ et $\pi_1(\theta|x)$ représentent les lois *a posteriori* de π_0 et π_1 .

Alors,

$$\begin{aligned}
 r(\pi) &= \lambda(x)r(\pi_0) + (1 - \lambda(x))r(\pi_1) \\
 \text{Avec } \lambda(x) &= \frac{(1 - \epsilon)m_{\pi_0(x)}}{m_{\pi}(x)}
 \end{aligned} \tag{4.11}$$

4.5.3 Cas où il y a plusieurs informations de mesure différentes provenant d'autres études

Soit π_0 la loi la plus "probable" de la loi *a priori* cherchée, π_1 correspond à une autre étude plus éloignée en sens d'"importance" et elle se décompose hiérarchiquement et appartient à la famille de lois qui s'écrit sous la forme :

$$\pi_1(\theta) = (1 - \epsilon_2)\pi_2(\theta) + \epsilon_2\pi_3(\theta)$$

Avec π_2 correspondant à la loi la plus proche de π_1 , et π_3 à celle plus éloignée.

D'où, en présence de "k" études "voisines", la loi *a priori* peut s'écrire sous la forme :

$$\pi(\theta) = (1 - \epsilon_1)\pi_0(\theta) + \sum_{k=1}^n \epsilon_1 \cdots \epsilon_k (1 - \epsilon_{k+1})\pi_{2k}(\theta) + \epsilon_1 \cdots \epsilon_{k+1}\pi_{2k+1}(\theta)$$

Alors, le cadre bayésien, nous obtenons le risque bayésien *a posteriori*, $r_\pi(x)$ par :

$$r_\pi(x) = \lambda_1 r_{\pi_0}(x) + \sum_{l=1}^k \left[\left(\prod_{i=1}^l (1 - \lambda_i) \right) \lambda_{l+1} r_{\pi_{2l}}(x) \right] + \prod_{j=1}^{l+1} (1 - \lambda_j) r_{\pi_{2l+1}}(x)$$

Preuve: Nous avons démontré dans [Proposition 4.5.1](#) que :

$$\pi(\theta|x) = \lambda_1(x) \cdot \pi_0(\theta|x) + [1 - \lambda_1(x)] \pi_1(\theta|x)$$

$$\text{Avec : } \lambda_1(x) = \frac{(1-\epsilon_1)m_{\pi_0}(x)}{m_\pi(x)}.$$

Et nous savons que :

$$r(\pi) = \int \theta \pi(\theta|x) d\theta$$

et que :

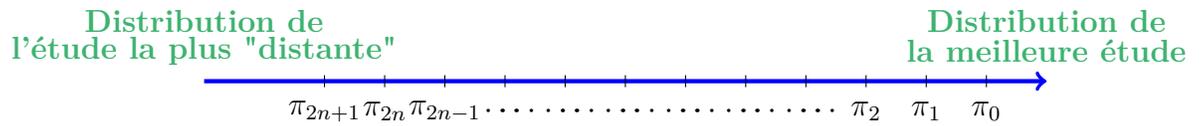
$$r(\pi) = \lambda_1(x)r(\pi_0) + (1 - \lambda_1(x))r(\pi_1) \quad 4.11$$

Donc pour n= 7 études nous avons :

$$\begin{aligned} r(\pi) &= \lambda_1(x)r(\pi_0) + (1 - \lambda_1(x))\lambda_2(x)r(\pi_2) + (1 - \lambda_1(x))(1 - \lambda_2(x))\lambda_3(x)r(\pi_4) \\ &\quad + (1 - \lambda_1(x))(1 - \lambda_2(x))(1 - \lambda_3(x))\lambda_4(x)r(\pi_6) + \\ &\quad (1 - \lambda_1(x))(1 - \lambda_2(x))(1 - \lambda_3(x))(1 - \lambda_4(x))r(\pi_7) \quad (4.12) \end{aligned}$$

Donc on peut l'écrire sous la forme :

$$r_{\pi}(x) = \lambda_1 r_{\pi_0}(x) + \sum_{l=1}^3 \left[\left(\prod_{i=1}^l (1 - \lambda_i) \right) \lambda_{l+1} r_{\pi_{2l}}(x) \right] + \prod_{j=1}^{3+1} (1 - \lambda_j) r_{\pi_{2l+1}}(x) \quad \square$$



4.5.4 Présentation de la pseudo-polyarthrite rhizomélique et sa relation avec l'artérite à cellules géantes

La polymyalgia rheumatica ou pseudo-polyarthrite rhizomélique est un trouble rhumatismal associé à des douleurs musculo-squelettiques modérées à sévères et à une rigidité du cou, des épaules et des hanches. La rigidité est plus visible le matin ou après une période d'inactivité. Ce trouble peut se développer rapidement dans certaines cas, mais pour la plupart des gens, la pseudo-polyarthrite rhizomélique se développe plus progressivement. La cause de la pseudo-polyarthrite rhizomélique n'est pas connue. Mais elle est associée à des anomalies du système immunitaire, des facteurs génétiques et un événement comme une infection, peut être un facteur déclenchant. Le fait que la pseudo-polyarthrite rhizomélique soit rare chez les personnes âgées de moins de 50 ans et devienne plus fréquente avec l'âge, suggère qu'elle peut être liée au processus de vieillissement.

La pseudo-polyarthrite rhizomélique se résout généralement en une à plusieurs années. Les symptômes de la pseudo-polyarthrite rhizomélique sont rapidement contrôlés par un traitement par corticostéroïdes, mais les symptômes récidivent si le traitement est interrompu trop tôt. Le traitement par les corticostéroïdes ne semble pas influencer la durée de la maladie.

La question qui nous intéresse dans cette partie c'est la relation entre la pseudo-polyarthrite rhizomélique et l'artérite à cellules géantes.

La raison pour laquelle la pseudo-polyarthrite rhizomélique et l'artérite à cellules géantes se produisent souvent ensemble n'est pas élucidée. Mais certaines personnes atteintes de pseudo-polyarthrite rhizomélique développent également une artérite à cellules géantes soit simultanément, soit après la disparition des symptômes musculo-squelettiques. D'autres personnes atteintes d'une maladie de Horton peuvent présenter également une pseudo-polyarthrite rhizomélique.

4.5.4.1 Calcul des paramètres des lois bêta dans la pseudo-polyarthrite rhizomélique

Chez les patients présentant une pseudo-polyarthrite rhizomélique, notons P_3 et P_4 , la probabilité du taux de non-rechute des patients dans le groupe CS et MTX, respectivement.

De même que P_1 et P_2 , P_3 et P_4 suivent des lois bêta. Pour calculer les paramètres correspondants, nous suivons les mêmes étapes présentées pour les patients atteints d'artérite à cellules géantes. Les mêmes questions Q1, Q2, Q5, Q6 sont soumises aux experts et leurs réponses sont placées dans le tableau 4.4.

| | Q1 | Q2 | Q5 | Q6 |
|-----------|------|------|-----|------|
| Expert 1 | 0.85 | 0.6 | 0.9 | 0.65 |
| Expert 2 | 0.8 | 0.65 | 0.9 | 0.6 |
| Expert 3 | 0.85 | 0.6 | 0.9 | 0.65 |
| Consensus | 0.85 | 0.6 | 0.9 | 0.65 |

Table 4.4 – Les réponses des experts aux questions Q1 à Q4 pour les études des patients atteints de PPR

Les réponses aux questions ci-dessus permettent de définir les paramètres des lois de P_3 et de P_4 :

$$P_3 \hookrightarrow \text{Beta}(3.77; 1.49)$$

$$P_4 \hookrightarrow \text{Beta}(4.15; 1.35)$$

4.5.5 Calcul des lois *a posteriori* avec la méthode ϵ -contamination

La loi *a priori* du taux de non-rechute dans le groupe CS correspond à :

$$\pi(P_{CS}) = (1 - \epsilon)P_1 + \epsilon P_3$$

et celle dans le groupe MTX correspond à :

$$\pi(P_{MTX}) = (1 - \epsilon)P_2 + \epsilon P_4$$

Avec P_1 et P_3 correspondant respectivement aux taux de non-rechute des CS seuls pour les patients atteints d'artérite à cellules géantes et pseudo-polyarthrite rhizomélique, et P_2 et P_4 correspondant respectivement aux taux de non-rechute du traitement par MTX et corticostéroïdes pour les patients atteints d'artérite à cellules géantes et pseudo-polyarthrite rhizomélique,

Donc

$$\pi(P_{CS}) = (1 - \epsilon)\beta(5.64, 2.99) + \epsilon\beta(3.77; 1.49) \quad (4.13)$$

et,

$$\pi(P_{MTX}) = (1 - \epsilon)\beta(3.77, 1.49) + \epsilon\beta(4.15; 1.35) \quad (4.14)$$

D'autre part nous considérons les données extraites de deux essais randomisés, le premier est l'essai randomisé de [Jover et al., 2001] qui montre que le traitement par méthotrexate associé aux corticostéroïdes est une alternative aux corticostéroïdes seuls chez les patients atteints d'artérite à cellules géantes et est plus efficace dans le contrôle de la maladie. Le deuxième essai randomisé en double aveugle est celui de [Caporali et al., 2004] pour comparer l'efficacité et la sécurité de la prednisone(corticostéroïdes) plus le méthotrexate vs la prednisone (corticostéroïde) seule chez les patients atteints de pseudo-polyarthrite rhizomélique. Donc les vraisemblances correspondent à des lois binomiales telles que : Dans l'essai de [Jover et al., 2001] pour le groupe recevant le corticostéroïde seul nous avons :

$$f(x_{Jover}|P_1) \propto B(19, \frac{3}{19})$$

Dans l'essai de [Caporali et al., 2004] pour le groupe recevant le corticostéroïde seul nous avons :

$$f(x_{Caporali}|P_3) \propto B(30, \frac{11}{30})$$

Dans l'essai de [Jover et al., 2001] pour le groupe recevant le méthotrexate associé aux CS nous avons :

$$f(x_{Jover}|P_2) \propto B(20, \frac{11}{20})$$

Dans l'essai de [Caporali et al., 2004] pour le groupe recevant le méthotrexate associé aux CS nous avons :

$$f(x_{Caporali}|P_4) \propto B(32, \frac{22}{32})$$

D'où la vraisemblance pour les patients recevant le corticostéroïde seul est :

$$f(x|P_{CS}) = (1 - \epsilon)B(19, \frac{3}{19}) + \epsilon B(30, \frac{11}{30}) \quad (4.15)$$

et celle pour les patients recevant le méthotrexate associé aux CS est :

$$f(x|P_{MTX}) = (1 - \epsilon)B(20, \frac{11}{20}) + \epsilon B(32, \frac{22}{32}) \quad (4.16)$$

4.5.6 Calcul de la différence entre les traitements

Nous cherchons à calculer $P(P_{MTX} - P_{CS} \geq -0.1|x_J, x_C, y_J, y_C)$, avec :
 x_J, x_C = valeur observée du taux de succès du CS dans les essais de [Jover et al., 2001] et [Caporali et al., 2004] respectivement.
 y_J, y_C = valeur observée du taux de succès du MTX dans les essais de [Jover et al., 2001] et [Caporali et al., 2004] respectivement.

Donc, en supposant que les études soient indépendantes et randomisées, nous avons :

$$\begin{aligned} & P(P_{MTX} - P_{CS} \geq -0.1|x_J, x_C, y_J, y_C), \\ &= \int_0^1 \left(\int_0^{P_{MTX}+0.1} \pi(P_{CS}|x_J, x_C) dP_{CS} \right) \pi(P_{MTX}|y_J, y_C) dP_{MTX} \\ &= \int_0^1 \left[\int_0^{P_{MTX}+0.1} \lambda(x_J, x_C) \pi(P_{CS}|x_J) + (1 - \lambda(x_J, x_C)) \pi(P_{CS}|x_C) dP_{CS} \right] \pi(P_{MTX}|y_J, y_C) dP_{MTX} \\ &= \int_0^1 \left[\int_0^{P_{MTX}+0.1} \lambda(x_J, x_C) \pi(P_{CS}|x_J) + (1 - \lambda(x_J, x_C)) \pi(P_{CS}|x_C) dP_{CS} \right] \\ &\quad \left[\mu(y_J, y_C) \pi(P_{MTX}|y_J) + (1 - \mu(y_J, y_C)) \pi(P_{MTX}|y_C) dP_{MTX} \right] \end{aligned}$$

Avec,

$$\begin{aligned} \lambda(x_J, x_C) &= (1 - \epsilon) \frac{m_{\pi_1}^{CS}}{m_{\pi}^{CS}} \\ &= (1 - \epsilon) \frac{m_{\pi_1}^{CS}}{(1-\epsilon)m_{\pi_1}^{CS} + \epsilon m_{\pi_3}^{CS}} \\ &= (1 - \epsilon) \frac{\int_0^1 f(x_J|P_{CS}) \pi_1(P_{CS}) dP_{CS}}{(1-\epsilon) \int_0^1 f(x_J|P_{CS}) \pi_1(P_{CS}) dP_{CS} + \epsilon \int_0^1 f(x_C|P_{CS}) \pi_3(P_{CS}) dP_{CS}} \end{aligned}$$

Et,

$$\begin{aligned} \mu(y_J, y_C) &= (1 - \epsilon) \frac{m_{\pi_2}^{MTX}}{m_{\pi}^{MTX}} \\ &= (1 - \epsilon) \frac{m_{\pi_2}^{MTX}}{(1-\epsilon)m_{\pi_2}^{MTX} + \epsilon m_{\pi_4}^{MTX}} \\ &= (1 - \epsilon) \frac{\int_0^1 f(y_J|P_{MTX}) \pi_2(P_{MTX}) dP_{MTX}}{(1-\epsilon) \int_0^1 f(y_J|P_{MTX}) \pi_2(P_{MTX}) dP_{MTX} + \epsilon \int_0^1 f(y_C|P_{MTX}) \pi_4(P_{MTX}) dP_{MTX}} \end{aligned}$$

Et,

$$\pi(p_1|x_1) \hookrightarrow \text{Beta}(x_1 + \alpha_1, n_1 - x_1 + \beta_1)$$

avec $p_1 \propto \text{Beta}(\alpha_1, \beta_1)$ et $f(x_1|p_1) \propto B(n_1, p_1)$ (Calcul présenté dans la section 13).

Donc,

$$\pi(P_{CS}|x_J) \propto \text{Beta}(5.64 + 3, 2.99 + 19 - 3) = \text{Beta}(8.64, 18.99)$$

$$\pi(P_{CS}|x_C) \propto \text{Beta}(3.77 + 11; 1.49 + 30 - 11) = \text{Beta}(14.77, 20.49)$$

$$\pi(P_{MTX}|y_J) \propto \text{Beta}(3.77 + 11; 1.49 + 20 - 11) = \text{Beta}(14.77, 10.49)$$

$$\pi(P_{MTX}|y_C) \propto \text{Beta}(4.15 + 22; 1.35 + 32 - 22) = \text{Beta}(26.15, 11.35)$$

D'où

$$\begin{aligned}
 & P(P_{MTX} - P_{CS} \geq -0.1 | x_J, x_C, y_J, y_C), \\
 &= \lambda(x_J, x_C) \mu(y_J, y_C) \int_0^1 (\int_0^{P_{MTX}+0.1} \text{Beta}(8.64; 18.99) dP_{CS}) \text{Beta}(14.77; 10.49) dP_{MTX} \\
 &+ \lambda(x_J, x_C) (1 - \mu(y_J, y_C)) \int_0^1 (\int_0^{P_{MTX}+0.1} \text{Beta}(8.64; 18.99) dP_{CS}) \text{Beta}(26.15; 11.35) dP_{MTX} \\
 &+ (1 - \lambda(x_J, x_C)) \mu(y_J, y_C) \int_0^1 (\int_0^{P_{MTX}+0.1} \text{Beta}(14.77; 20.49) dP_{CS}) \text{Beta}(14.77; 10.49) dP_{MTX} \\
 &+ (1 - \lambda(x_J, x_C)) (1 - \mu(y_J, y_C)) \int_0^1 (\int_0^{P_{MTX}+0.1} \text{Beta}(14.77; 20.49) dP_{CS}) \text{Beta}(26.15; 11.35) dP_{MTX}
 \end{aligned}$$

Avec :

$$\begin{aligned}
 & \int_0^1 (\int_0^{P_{MTX}+0.1} \text{Beta}(c; d) dP_{CS}) \text{Beta}(a; b) dP_{MTX} \\
 &= \int_0^1 (\int_0^{P_{MTX}+0.1} \frac{1}{B(c; d)} y^{c-1} (1-y)^{d-1}) \frac{1}{B(a; b)} x^{a-1} (1-x)^{b-1} dP_{CS} dP_{MTX} \\
 &= \frac{1}{B(c; d) B(a; b)} \int_0^1 (\int_0^{P_{MTX}+0.1} y^{c-1} (1-y)^{d-1} dP_{CS}) x^{a-1} (1-x)^{b-1} dP_{MTX}
 \end{aligned}$$

Donc, Pour $\epsilon = 0.1$ nous obtenons :

$$\lambda(x_J, x_C) = 0.1 \text{ et } \mu(y_J, y_C) = 0.8$$

Alors,

$$P(P_{MTX} - P_{CS} \geq -0.1 | x_{JC}) = 0.3102 \approx 31\%$$

Pour $\epsilon = 0.2$ nous avons :

$$\lambda(x_J, x_C) = 0.05 \text{ et } \mu(y_J, y_C) = 0.68$$

Alors,

$$P(P_{MTX} - P_{CS} \geq -0.1 | x_J, x_C, y_J, y_C) = 0.348 \approx 35\%$$

Pour $\epsilon = 0.3$ nous avons :

$$\lambda(x_J, x_C) = 0.028 \text{ et } \mu(y_J, y_C) = 0.523$$

Alors,

$$P(P_{MTX} - P_{CS} \geq -0.1 | x_J, x_C, y_J, y_C) = 0.376 \approx 38\%$$

4.5.7 Interprétation

Le résultat de l'essai de [Jover et al., 2001] et [Caporali et al., 2004] sont significatifs, $P=0.004$ et $P=0.02$ respectivement. Mais les résultats trouvés sont compris entre 30% et 40%, ce qui montre qu'il y a un doute sur la non-infériorité du MTX associé aux CS par rapport aux CS seuls.

Dans ce cas, les méthodes bayésiennes ne semblent pas apporter plus d'informations, si ce n'est le fait que les cliniciens peuvent porter un jugement sur les taux obtenus (31%, 35%, 38%) et leur acceptabilité en pratique clinique.

Conclusion

Nous avons développé dans le premier chapitre, une large revue de la littérature permettant de guider à la prise de décision dans le cas des maladies rares. Cette recherche sur les méthodes associées aux maladies rares comporte plusieurs limites. Tout d'abord, notre recherche sur la littérature a porté sur tous les articles liés aux maladies rares. En raison du grand nombre de maladies rares individuelles, nous n'avons pas pu rechercher des applications de méthodes innovantes liées à chaque maladie spécifique.

De plus, notre recherche avait pour but de donner un aperçu général des méthodes non traditionnelles qui ont été proposées ou appliquées à l'étude des maladies rares. Si d'autres méthodes non traditionnelles existent qui pourraient être applicables aux maladies rares, mais qui n'ont pas encore été discutées dans une publication dans les bases de données que nous avons recherchées, nous n'avons pas pu les identifier.

En plus, notre revue de la littérature avait pour but de mieux faire connaître la disponibilité et l'utilisation des méthodes novatrices pour la gestion des maladies rares et des essais cliniques de petite taille et ne visait pas à fournir tous les détails de calcul de ces méthodes, détails qui peuvent être trouvés dans les références citées.

Sur la base des articles étudiés, nous pensons qu'il n'y a que peu de prise de conscience sur les possibilités statistiques pour améliorer les preuves dans le cas des études sur petits échantillons. En dehors de cela, nous avons observé un manque d'étude systématique de la conception et des méthodes statistiques pour les études sur les maladies rares, de sorte que d'autres recherches seraient souhaitables. Cependant, les recommandations générales sont difficiles dans le domaine des maladies rares, car les situations et les tailles d'échantillons varient énormément.

Des approches novatrices peuvent permettre aux chercheurs d'aborder des questions auxquelles on ne peut répondre avec des approches conventionnelles.

Plusieurs des approches présentées dans le premier chapitre peuvent être utilisées en combinaison avec la méthode bayésienne ou bien avec d'autres approches.

Nous proposons dans le deuxième chapitre l'utilisation d'une approche bayésienne qui peut augmenter la robustesse de l'information à partir de petits essais et peut être utilisée pour prendre une décision avec des tels essais. Cette approche utilise toutes les données disponibles, de l'essai et d'autres sources pour calculer les probabilités qu'un traitement particulier soit efficace. Ces probabilités peuvent ensuite être appliquées à la pratique clinique pour proposer une décision ou bien avoir une meilleure compréhension de la problématique.

Nous présentons dans le chapitre trois comment éliciter les avis des experts. Souvent les chercheurs sollicitent l'avis de plus d'un expert quand les décisions sont prises en présence d'incertitude. Il peut alors être nécessaire ou souhaitable d'obtenir une distribution unique qui synthétise les croyances de plusieurs experts. Dans ce cadre, nous avons présenté deux approches pour combiner les différentes distributions *a priori* en un seul consensus d'*a priori*. Ces deux approches sont l'agrégation mathématique ou comportementale au sein de chaque approche globale. Mais le débat sur l'agrégation mathématique versus l'agrégation comportementale n'est pas encore résolu, car il n'y a pas d'évidences pour favoriser une méthode plutôt qu'une autre et chaque approche a ses avantages et ses inconvénients. Les approches mathématiques sont faciles à utiliser. Mais elles n'ont pas de base normative pour choisir des poids. En outre, plusieurs résultats montrent qu'aucune fonction d'agrégation ne peut satisfaire toutes les propriétés souhaitées d'un pool d'opinions, à moins que l'opinion commune ne converge vers en une seule opinion individuelle, ce qui implique effectivement un leader d'opinion.

Nous avons présenté dans le dernier chapitre trois méthodes pour concevoir et interpréter les essais cliniques menés avec des tailles d'échantillons limitées dans le cas des patients atteints d'artérite à cellules géantes ou maladie de Horton.

Nous présentons dans la première méthode une approche bayésienne pour la conduite d'essais de maladies rares. Le but de cette approche est de vérifier l'hypothèse de non-infériorité de l'essai qui repose sur l'hypothèse que le méthotrexate (MTX) associé aux corticostéroïdes (CS) est plus efficace que les corticostéroïdes (CS) seuls, où les réponses des patients sont classées comme un succès ou un échec.

L'information *a priori* sur les paramètres de ce modèle bayésien est résumée par les lois *a priori*. Nous présentons alors, l'élicitation des avis des experts qui consiste à obtenir des réponses à des questions précises afin de calculer les paramètres des lois *a priori* du taux de non-rechute dans le groupe des corticostéroïdes seuls et la loi du log odds-ratio qui représente l'efficacité du méthotrexate associé aux corticostéroïdes par rapport aux

corticostéroïdes seuls.

Une fois que les lois *a priori* ont été établies, nous considérons pour chaque échantillon, les valeurs possibles de succès et d'échec dans le groupe MTX associé aux CS et dans le groupe CS seuls. Les distributions *a posteriori* correspondantes peuvent alors être calculées.

Ensuite, nous calculons les probabilités de non-infériorité, telles que le taux de patients ne présentant aucune rechute dans le groupe MTX soit supérieur à celui du groupe CS moins un seuil de 10%, sachant les données de trois essais randomisés "voisins".

Donc, les probabilités de non-infériorité obtenues pour les trois études sont supérieures à 60%, ce qui reste en faveur du méthotrexate par rapport aux corticostéroïdes seuls.

Nous proposons dans la deuxième méthode, une approche bayésienne pour vérifier l'hypothèse de non-infériorité de l'efficacité du méthotrexate associé aux corticostéroïdes par rapport aux corticostéroïdes seuls. Nous cherchons à calculer les paramètres des lois *a priori* du taux de non-rechute dans le groupe des corticostéroïdes seuls et celui du groupe méthotrexate. Pour trouver ces paramètres nous présentons un deuxième questionnaire aux experts, puis à partir des réponses obtenues au premier questionnaire et celles du deuxième questionnaire nous avons obtenu les paramètres des deux lois *a priori* correspondantes. Ensuite, nous calculons les probabilités de non-infériorité, telle que le taux de patients ne représentant aucune rechute dans le groupe MTX soit supérieur à celui du groupe CS moins un seuil de 10%, sachant les données de trois essais randomisés "voisins", cette probabilité est présentée par le calcul de la probabilité de la différence de pourcentage de succès entre les deux groupes de traitement.

Les résultats obtenus par cette méthode sont concordants avec ceux obtenus sur l'efficacité du MTX associé aux CS par rapport à ceux du groupe CS présentés dans les études.

Enfin, nous proposons dans ce chapitre une approche bayésienne non utilisée auparavant dans la littérature pour la conduite des essais de maladies rares comparant un traitement expérimental à un traitement contrôle, la méthode ϵ -contamination. Cette méthode se base sur le principe de contaminer une loi *a priori*, pas entièrement satisfaisante par une série de lois provenant des informations d'autres études portant sur une même pathologie, même traitement ou concernant une population analogue.

Les probabilités de non-infériorité sont comprises entre 30% et 40%, les résultats obtenus par cette dernière méthode sont inférieurs naturellement à ceux obtenus par les méthodes utilisées sur chaque étude séparément.

Dans ce cas, les méthodes bayésiennes semblent apporter une autre information condui-

sant les praticiens à l'utiliser afin de prendre une décision notamment dans le cas de la maladie de Horton.

| Intérêt | Originalité | Limites |
|--|--|---|
| <ul style="list-style-type: none"> • Proposition et comparaison des méthodes statistiques dans le cas des maladies rares. • Intérêt des méthodes bayésiennes. • Elicitation des avis des experts. • Etude de non-infériorité pour comparer l'efficacité d'un nouveau traitement avec le traitement de référence. | <ul style="list-style-type: none"> • Notre revue de la littérature est un support destiné aux professionnels impliqués à l'étude des maladies. • L'élicitation des avis des experts dans la méthode de différence de deux lois bêta. • La méthode ϵ-contamination jamais été utilisée dans le contexte des maladies rares et en recherche médicale. • Fournir une aide à la décision thérapeutique à partir des 3 méthodes présentées précédemment | <ul style="list-style-type: none"> • Part de subjectivité dans l'estimation des lois a priori dans le contexte d'élicitation des préférences. • Pas de comparaison possible des 3 méthodes: <ul style="list-style-type: none"> - Approche étude par étude (2 premières méthodes) - Approche par mélange d'étude (méthode ϵ-contamination) |

Figure 4.7 – Tableau représentatif de l'intérêt et l'originalité de la thèse ainsi que les limites

Annexe

4.6 Annexe du chapitre 1

Table 4.5 – *Présentation des méthodes utilisées dans le cas des maladies rares*

| Méthodes | Description résumée | Avantages | Difficultés |
|-----------------------------|--|---|---|
| Modèle de démarrage retardé | Les patients sont randomisés pour recevoir le même traitement actif, mais en commençant à des moments différents | <ul style="list-style-type: none">– Améliorer la compréhension sur l'efficacité des traitements– Modification de la maladie et amélioration symptomatique dans le groupe de traitement pendant la phase 1 d'essai. | <ul style="list-style-type: none">– Problème des données manquantes,– Difficulté du choix de la durée du retard pour le démarrage de la phase 2,– Choix des tests statistiques à la fin de la phase 2– Effet "carry-over" entre la première et la deuxième phase ne peut pas être exclu. |

Table 4.5 – *Présentation des méthodes utilisées dans le cas des maladies rares*

| Méthodes | Description résumée | Avantages | Difficultés |
|-------------------|---|---|---|
| Essais cross-over | <p>Les patients sont assignés au hasard à un des traitements, suivi d'une période "washout", puis reçoivent le traitement alternatif (il peut s'agir de la comparaison de 2 ou de plus de 2 traitements). D'autres patients sont randomisés dans le sens inverse. Les patients peuvent être affectés au hasard à chaque traitement à de multiples points dans le temps.</p> | <ul style="list-style-type: none"> – Garantit que le sujet aura les 2 traitements donc améliore le recrutement. – Chaque participant est seul propre contrôle minimisant la variance et l'effet des variables des confiances. – $\frac{N}{2}(1 - \rho)$ le nombre des patients que celui dans un groupe parallèle | <ul style="list-style-type: none"> – Risque de perdus de vue. – Demande une réponse rapide. – Approprié que pour les maladies classiques avec effet rapide du traitement. – Risque d'un effet période. – Risque d'un effet carry-over. |

Table 4.5 – Présentation des méthodes utilisées dans le cas des maladies rares

| Méthodes | Description résumée | Avantages | Difficultés |
|----------------------|---|---|--|
| Essai de type N-of-1 | C'est le principe du cross-over avec un seul participant | <ul style="list-style-type: none"> – Garantit que le sujet aura les 2 (ou n) traitements donc améliore le recrutement. – Le participant est son propre témoin minimisant la variance et l'effet des variables des confusions. | <ul style="list-style-type: none"> – Demande une réponse rapide. – Approprié que pour les maladies classiques. – Risque d'un effet période. – Risque d'un effet carry-over. – La méta-analyse des essais N-of-1 est limitée par l'hétérogénéité des essais entre individus. |
| Plan Factoriel | Fourmillent des réponses à plusieurs questions dans la même population d'étude et permettent aux chercheurs d'évaluer plus d'une intervention en une seule expérience | <ul style="list-style-type: none"> – Réduit le nombre total de patients requis pour répondre à toutes les questions. – Permet de tester les interactions entre les interventions comparées. | <ul style="list-style-type: none"> – Peut avoir un protocole d'étude plus complexe : critères d'évaluation et suivi des patients plus rigoureux. – Risque de manque de puissance pour tester les interactions. |

Table 4.5 – Présentation des méthodes utilisées dans le cas des maladies rares

| Méthodes | Description résumée | Avantages | Difficultés |
|------------------------------|--|--|---|
| Design de retrait aléatoire | Utilisé pour évaluer la continuation du traitement chez les patients qui répondent au traitement. Tous les patients reçoivent d'abord le traitement testé, les intervenants sont randomisés pour poursuivre le traitement ou pour recevoir un placebo ou le traitement de référence. | <ul style="list-style-type: none"> – Réduit le temps sous placebo puisque seuls les patients qui répondent sont randomisés au placebo ou le traitement de référence.. – Peut évaluer si le traitement doit être poursuivi ou peut être arrêté. | <ul style="list-style-type: none"> – Utilisation dans les maladies chroniques, ne convient pas aux maladies imprévisibles (par exemple une rémission spontanée) ou ceux à évolution lente. – L'effet du traitement est surestimé puisque seuls les répondants sont inclus (et comparés au placebo ou le traitement de référence.) |
| Modèle d'échappement précoce | Les patients sont retirés s'ils satisfont à un critère de défaillance <i>a priori</i> | <ul style="list-style-type: none"> – Réduit le temps sous placebo ou l'échec du traitement. – Analyser le taux d'échec, afin de minimiser l'exposition à un traitement inefficace | <ul style="list-style-type: none"> – Difficile de définir un résultat binaire échec/réussite. – Seulement l'efficacité à court terme a été évaluée. – Perte de pouvoir si un nombre important de patients "échappent" |

Table 4.5 – *Présentation des méthodes utilisées dans le cas des maladies rares*

| Méthodes | Description résumée | Avantages | Difficultés |
|---|--|---|---|
| Conception randomisée de la phase placebo | <p>Tous les patients reçoivent le traitement expérimental testé - mais ont des durées variables sous placebo. Suppose qu'une réponse se produira quelque temps après que le traitement efficace est donné, de sorte que les patients qui commencent le traitement plus tôt devraient, en moyenne, répondre plus tôt.</p> | <ul style="list-style-type: none"> - Peut être utilisé pour des thérapies modifiant la maladie, dans des maladies à évolution rapide et défavorable. - Tous les patients reçoivent un traitement actif. | <ul style="list-style-type: none"> - La durée variable de la période placebo réduit la puissance statistique. - Capacité limitée d'estimer la taille de l'effet de traitement |

Table 4.5 – *Présentation des méthodes utilisées dans le cas des maladies rares*

| Méthodes | Description résumée | Avantages | Difficultés |
|---|--|--|---|
| Design de randomisation adapté à la réponse des patients précédents (response-adaptive randomisation) | Ce design maximise le nombre de participants affectés au traitement le plus efficace en minimisant le recrutement global de l'étude. | Réduit le nombre de sujets ayant le traitement le moins efficace | <ul style="list-style-type: none"> – Risque de biais de sélection dû à l'absence d'aveugle. – Biais dû au changement temporel des caractéristiques du patient ou du traitement. – La complexité du protocole augmente la difficulté pour avoir un accord éclairé |

Table 4.5 – Présentation des méthodes utilisées dans le cas des maladies rares

| Méthodes | Description résumée | Avantages | Difficultés |
|--|---|--|--|
| Schéma de randomisation adaptative fondé sur les covariables (Covariate-adaptive approach) | C'est un schéma de randomisation qui repose sur une stratification selon les covariables initiales et qui alloue le traitement pour assurer un équilibre dans chaque strate. Dans ce design un nouveau participant est séquentiellement attribué à un groupe de traitement particulier en prenant en compte les variables spécifiques et les affectations précédentes des participants. La randomisation adaptative sur covariable utilise la méthode de la minimisation en évaluant le déséquilibre de la taille de l'échantillon entre plusieurs variables. | <ul style="list-style-type: none"> – Permet l'examen des allocations des groupes des participants précédents pour prendre une décision au cas par cas sur la base de l'affectation du groupe pour chaque personne enrôlée dans l'étude. – Produit moins de déséquilibre que les autres méthodes de randomisation classiques. – Peut être utilisé avec succès pour équilibrer des covariables importantes entre les groupes contrôles et expérimentales. | <ul style="list-style-type: none"> – L'équilibre des variables entre les groupes devient difficile lorsque le nombre de blocs approche de la moitié de la taille de l'échantillon. – Les affectations de traitement deviennent parfois prévisibles. – Processus de calcul compliqué, augmente la charge administrative, ce qui limite son utilisation dans la pratique. |

Table 4.5 – *Présentation des méthodes utilisées dans le cas des maladies rares*

| Méthodes | Description résumée | Avantages | Difficultés |
|--------------------------------|---|---|--|
| Seamless adaptive trial design | C'est un modèle d'essai qui combine deux essais séparés dans un seul essai et utilise dans l'analyse finale les données provenant des patients inscrits avant et après l'adaptation. | <ul style="list-style-type: none"> – Peut être utilisé pour combiner la phase II et la phase III des études d'un même médicament. | <ul style="list-style-type: none"> – Moins de contrôle sur le taux d'erreur globale de type I. – Effectuer une analyse de puissance de calcul de taille de l'échantillon peut être problématique. – Exécution d'une analyse combinée des objectifs de l'étude/paramètres qui sont distincts à différentes phases. |
| An internal pilot design | C'est une étude adaptative dans laquelle la 1ere phase est dite phase pilote. Les investigateurs peuvent estimer : <ul style="list-style-type: none"> – Le taux global d'événements. – La variance inter sujet. | <ul style="list-style-type: none"> – Réduit la perte des participants éligibles, car l'étude est pilote. – Fournit une mesure précise de la variance pour un échantillon approprié de l'étude principale. | Risque d'augmentation de l'erreur de type I |

Table 4.5 – Présentation des méthodes utilisées dans le cas des maladies rares

| Méthodes | Description résumée | Avantages | Difficultés |
|-------------------|--|--|--|
| Design séquentiel | <p>Design séquentiel similaire à un RCT traditionnel, mais permet des analyses intermédiaires répétées pour guider une fin d'étude prématurée lorsque des préoccupations de sécurité, la futilité, l'efficacité, ou une combinaison de ces facteurs est démontrée. Contrairement aux essais traditionnels, le nombre d'individus à inclure dans l'essai n'est pas connu à l'initiation. On peut faire l'analyse après chaque patient inclus (séquentiel continu) ou après un nombre fixe ou variable de patients (séquentiel groupé)</p> | <p>Permet de terminer l'étude précocement.</p> | <ul style="list-style-type: none"> – Le résultat des traitements doit être rapide après l'inclusion. – Risque d'augmentation de l'erreur de type I. – Complexité de déterminer le temps de la première analyse, les méthodes pour analyser les traitements multiples, le nombre optimal des analyses intermédiaires, et la méthode pour calculer la puissance de l'étude. |

Table 4.5 – *Présentation des méthodes utilisées dans le cas des maladies rares*

| Méthodes | Description résumée | Avantages | Difficultés |
|--|--|---|---|
| Design de type écarter le perdant, retenir le gagnant. | C'est un design qui permet d'abandonner les groupes de traitements d'efficacité inférieure et permet aussi d'ajouter des bras supplémentaires. | <ul style="list-style-type: none"> – Utile dans la phase II de développement clinique surtout quand il y a des incertitudes concernant les doses. – Réduit le nombre de patients recevant un traitement moins efficace. – Pourrait améliorer le recrutement des patients grâce à une meilleure satisfaction. | <ul style="list-style-type: none"> – La taille d'échantillon inégale réduit la puissance. – Dans certaines situations, le nombre de patients qui ont effectivement reçu un des traitements est très faible. – Les groupes soumis aux doses qui sont abandonnées peuvent contenir des informations précieuses pour la connaissance de la réponse au traitement. |

Table 4.5 – *Présentation des méthodes utilisées dans le cas des maladies rares*

| Méthodes | Description résumée | Avantages | Difficultés |
|------------------------------|---|--|---|
| Recherche de dose adaptative | Un design de type recherche de dose adaptative est souvent utilisé dans le développement clinique en phase précoce pour identifier la dose efficace minimale et la dose maximale tolérable, qui détermine la posologie à retenir pour les essais cliniques de phase suivante. | <ul style="list-style-type: none"> – Abandonner précocement les groupes recevant les doses les plus basses. – Modifie/ajoute des groupes supplémentaires de doses. | <ul style="list-style-type: none"> – Sélectionner la dose initiale. – Sélectionner la gamme de doses à l'étude. – Obtenir une signification statistique de puissance souhaitée avec un nombre limité de sujets. – Établir les critères de sélection et les règles de décision. – Estimer la probabilité de la réalisation de la dose optimale. – Gérer la chaîne d'approvisionnement des fournitures cliniques peut être problématique. |

Table 4.5 – *Présentation des méthodes utilisées dans le cas des maladies rares*

| Méthodes | Description résumée | Avantages | Difficultés |
|---|---|--|--|
| Design de réestimation de la taille d'échantillon | Ajustement de la taille de l'échantillon ou réestimation basée sur les données observées provisoirement | <ul style="list-style-type: none"> – Donne la taille convenable de l'essai pour démontrer l'efficacité. – Permet d'incorporer de nouvelles informations interne ou externe en cours d'essai. | <ul style="list-style-type: none"> – Difficulté de savoir s'il faut commencer par une taille d'échantillon large ou petite. – Les méthodes standard de réestimation de taille d'échantillon basées sur la différence observée avec un nombre limité de sujets peuvent être biaisées. |

Table 4.5 – *Présentation des méthodes utilisées dans le cas des maladies rares*

| Méthodes | Description résumée | Avantages | Difficultés |
|---------------------------|--|--|---|
| Biomarker-adaptive design | Basé sur les réponses des biomarqueurs tels que les marqueurs génomiques pour évaluer l'effet du traitement. Elle implique la qualification des biomarqueurs, design de sélection optimale | <ul style="list-style-type: none"> – Sélectionner la population convenable de malades (processus d'enrichissement pour la sélection d'une meilleure population cible de patients). – Identifier le stade de la maladie selon sa nature – Détecter précocement la maladie. – Aider au développement de la médecine personnalisée. | <ul style="list-style-type: none"> – Sélection de la gamme de doses à l'étude. – Besoin d'investir dans le séquençage du génome. – Décalage entre l'identification des gènes qui sont associés à des résultats cliniques et l'établissement d'un modèle de prédiction entre les gènes et les résultats cliniques pertinents. |

Table 4.5 – *Présentation des méthodes utilisées dans le cas des maladies rares*

| Méthodes | Description résumée | Avantages | Difficultés |
|----------------------------|---|---|---|
| Adaptive-hypotheses design | <p>C'est un design qui permet des changements dans les hypothèses basées selon les résultats d'une analyse intermédiaire :</p> <ol style="list-style-type: none"> 1. Échanger une hypothèse de supériorité pour une hypothèse de non-infériorité. 2. Changer le critère d'évaluation principal. | Assurer le succès des essais cliniques pour identifier les meilleurs avantages cliniques possibles pour les patients qui entrent dans les essais cliniques. | <ul style="list-style-type: none"> – Difficultés dans la sélection de la marge de non-infériorité. – Calcul de la taille d'échantillon. |
| Scores de propension | Valeur quantitative qui résume les caractéristiques initiales des sujets et sa liaison à l'indication du traitement pour constituer a posteriori des groupes comparables et ne différant que par le traitement | Réduire la confusion en résumant tous les facteurs de confusion potentiels en un seul score. | Le contrôle statistique de la confusion peut ne pas être toujours possible |

Table 4.5 – *Présentation des méthodes utilisées dans le cas des maladies rares*

| Méthodes | Description résumée | Avantages | Difficultés |
|--|--|---|---|
| Utilisation des biomarqueurs comme résultats de substitution | Peut être mesurée avant que les patients ne soient perdus de vue pour les paramètres cliniques difficiles | <ul style="list-style-type: none"> – Diminution des besoins en taille d'échantillon. – Effet net de l'intervention sur la mesure de l'efficacité clinique | <ul style="list-style-type: none"> – La validation des biomarqueurs comme des bons substituts des résultats cliniques peut être difficile |
| Planifier une méta-analyse prospective | Les petites études individuelles peuvent ne pas fournir de preuves définitives sur une question, mais peuvent être combinées en une méta-analyse prospective pour produire suffisamment de puissance | <ul style="list-style-type: none"> – Des risques excessifs de α et β peuvent être évités – Ils fournissent une assurance suffisante que les résultats d'une étude contribuent à accroître les connaissances. | <ul style="list-style-type: none"> – Trop peu de sujets pour pouvoir conclure positivement – Difficultés à synthétiser les résultats peuvent empêcher le calcul des effets de traitement valides. |

Table 4.5 – Présentation des méthodes utilisées dans le cas des maladies rares

| Méthodes | Description résumée | Avantages | Difficultés |
|----------------|---|---|---|
| Cadre Bayésien | <p>Méthodes bayésiennes pour l'estimation de l'effet du traitement.</p> <p>Avoir une estimation préalable de la taille de l'effet et mise à jour de cette estimation lorsque de nouvelles données s'accumulent.</p> | <ul style="list-style-type: none"> – Inclut l'information provenant d'autres études de types variés. – Permet le calcul de l'estimation ponctuelle, de l'intervalle de crédibilité et des distributions des densités pour les effets des traitements. | <ul style="list-style-type: none"> – Loi <i>a priori</i> peut être incertaine, subjective ou limitée par l'hétérogénéité de la variable de jugement. – Complexité de calcul de l'analyse. |

4.7 Annexe du chapitre 2

Algorithme pour l'exemple 2.3.1

```
calcPosteriorForProportion <- fonction(successes, total, a, b){  
Maximum de vraisemblance :  
library(bbmle)  
mydata<- c(12)  
myfunc <- fonction(size,prob) -sum(dbinom(mydata,size,prob,log=TRUE))  
myfunc  
mle2(myfunc, start=list(prob=0.4), data=list(size=20))  
Plot the prior, likelihood and posterior :  
likelihooda = successes + 1; likelihoodb = total - successes + 1  
posteriora = a + successes; posteriorb = b + total - successes  
theta = seq(0.005, 0.995, length = 500)  
theta1 = seq(0.05, 0.95, length = 500)  
prior = dbeta(theta, a, b)  
likelihood = dbeta(theta, likelihooda, likelihoodb)  
posterior = dbeta(theta, posteriora, posteriorb)  
m = max(c(prior, likelihood, posterior))  
plot(theta, posterior, type = "l", ylab = "Densite", lty = 2, lwd = 3,  
ylim = c(0, m), col = "blue")  
lines(theta1, likelihood, lty = 1, lwd = 3, col = "green")  
lines(theta, prior, lty = 3, lwd = 3, col = "red")  
legend(x = 0.7, y = m, c("APosteriori", "vraisemblance", "APriori"), lty = c(2, 1, 2),  
lwd = c(2, 3, 3), col = c("blue", "green", "red"), cex = 0.7)
```

4.7.1 Interchangeabilité et mise en place des modèles hiérarchiques

Considérons un ensemble d'expériences $j=1, \dots, N$, dans lequel l'expérience j à des données y_j et des paramètres θ_j avec fonction de vraisemblance $p(y_j|\theta_j)$. Certains des paramètres de différentes expériences peuvent se superposer ; par exemple chaque vecteur y_j peut être un échantillon des observations d'une distribution normale avec moyenne μ_j

et variance commune σ^2 , dans ce cas $\theta_j = (\mu_j, \sigma^2)$.

Afin de créer un modèle de probabilité joint pour tous les paramètres θ , on utilise l'idée d'interchangeabilité.

L'analyse statistique commence par l'hypothèse que les valeurs du vecteur $y = (y_1, y_2, \dots, y_n)$ peuvent être considérées comme échangeables, cela signifie qu'on exprime l'incertitude comme une fonction de densité de probabilité jointe invariante pour les permutations des indices. Un modèle non-échangeable peut être approprié si l'information pertinente pour le résultat a été transmise par l'unité d'indices plutôt que par les variables explicatives.

Si aucune information, autre que les données y , n'est disponible pour distinguer n'importe quel θ_i des autres, et aucun tri ou regroupement des paramètres peut être fait, il faut supposer une symétrie entre les paramètres dans leurs distributions a priori. Cette symétrie est représentée de manière probabiliste par interchangeabilité, les paramètres $(\theta_1, \dots, \theta_N)$ sont échangeables dans leurs distributions jointes si $p((\theta_1, \dots, \theta_N))$ est invariante pour les permutations des indices $(1, \dots, N)$. En pratique, l'ignorance implique l'interchangeabilité. Généralement, moins on a de connaissance sur le problème, plus on est sûr du choix de l'échangeabilité, mais ça ne veut pas dire de limiter notre connaissance du problème, car si on prend l'exemple d'un jet de dé : il faut d'abord attribuer des probabilités égales à toutes les six faces, mais si l'on étudie les mesures et le poids du dé attentivement, nous pourrions éventuellement remarquer des imperfections, ce qui pourrait nous faire privilégier un résultat plus qu'un autre et ainsi éliminer l'interchangeabilité entre les six résultats.

Le modèle le plus simple d'une loi interchangeable considère chaque paramètre θ_j comme un échantillon indépendant de la distribution a priori (ou population) gouverné par un vecteur de paramètres inconnus ψ , tel que,

$$p(\theta, \psi) = \prod_{j=1}^N p(\theta_j | \psi).$$

Définition 4.7.1. Une séquence de variables aléatoires $(X_i)_{i=1}^n$ est dite "finiment" échangeable si les distributions de n'importe quelle permutation est la même que toute autre permutation :

$$P\left(\bigcap_{i=1}^n X_{\pi(i)}\right) = P\left(\bigcap_{i=1}^n X_{(i)}\right)$$

Pour toutes fonctions de permutations $\pi(\cdot)$.

On peut étendre cette définition d'interchangeabilité à des séquences infinies de variables.

Même si les observations ne sont pas complètement échangeables, mais partiellement ou conditionnellement échangeables, on a deux cas :

D'une part, si les observations peuvent être groupées, on doit faire un modèle hiérarchique, où chaque groupe a son propre sous-modèle mais les propriétés des groupes sont inconnues. Si on suppose que les propriétés des groupes sont échangeables, on peut utiliser une distribution a priori commune pour les propriétés du groupe.

D'autres part, si y_i a des informations additionnelles x_i , tel que y_i n'est pas échangeable mais les covariables (y_i, x_i) sont échangeables, dans ce cas on peut faire un modèle joint pour (y_i, x_i) ou un modèle conditionnel $y_i|x_i$.

En général, le modèle usuel d'échangeabilité avec covariable est à travers des indépendances conditionnelles :

$$p(\theta_1, \dots, \theta_N, x_1, \dots, x_N) = \int \left(\prod_{j=1}^N p(\theta_j | \psi, x_j) \right) p(\psi | x) f(x) d\psi$$

avec $x = (x_1, \dots, x_N)$.

Dans cette façon, les modèles d'interchangeabilité deviennent applicables universellement, car toute information disponibles pour distinguer les différents groupes doit être codées dans les variables x et y .

Dans presque toutes les applications statistiques, c'est naturel de s'opposer à l'interchangeabilité dans le cas où les groupes diffèrent, par exemple si on a des expériences qui ont été effectuées à des temps différents, sur différents individus, et dans différents laboratoires. De telles informations ne valident pas la non interchangeabilité, car si les expériences diffèrent ceci implique que les θ_i sont différents, mais il peut être acceptable de les considérer comme s'ils ont été tirés d'une distribution commune. En effet, sans information disponible pour les distinguer, on n'a pas d'autre choix logique que le modèle θ_i échangeable.

4.7.2 Preuve du Théorème 2.61

Théorème 2.61

Un estimateur minimisant le risque bayésien $r(\pi, \delta)$ peut être obtenu en choisissant pour $x \in X$, la valeur $\delta(x)$ qui minimise le coût a posteriori, $\rho(\pi, \delta|x)$ puisque :

$$r(\pi, \delta) = \int_X \rho(\pi, \delta|x) m(x) dx$$

Preuve :

On a par définition

$$\begin{aligned}r(\pi, \delta) &= E^\pi[R(\theta, \delta)] \\&= \int_{\Theta} \int_X (R(\theta, \delta) \pi(\theta)) d\theta dx \\&= \left(\int_{\Theta} \int_X L(\theta, \delta(x)) f(x|\theta) dx \right) \pi(\theta) d\theta \\&= \int_X \int_{\Theta} L(\theta, \delta(x)) f(x|\theta) \pi(\theta) d\theta dx \quad (\text{d'après le théorème de Fubini}) \\&= \int_X \int_{\Theta} L(\theta, \delta(x)) \pi(\theta|x) d\theta m(x) dx \quad (\text{d'après le théorème de Bayes, avec } m(x) \text{ la loi marginale}) \\&= \int_X \rho(\pi, \delta|x) m(x) dx.\end{aligned}$$

4.8 Chapitre 4

4.9 Algorithme de calcul

Algorithme du calcul des paramètres de la loi bêta :

Fonction pour calculer les paramètres de la loi beta.

mode correspond à la réponse à Q1, et intvl celle à Q2

```
prior_beta <- fonction(mode, intvl) {  
  
  beta_calc <- fonction(a, mode, intvl){  
    b = ((a - 1)/mode) - (a-2)  
    x = qbeta(0.25, shape1 = a, shape2 = b, lower.tail=TRUE) - intvl  
    return(x)  
  
    lim1 = as.double(0.001)  
    lim2 = as.double(0.999)  
  
    fval = vector(mode="numeric", length=2)  
    fval[1] = beta_calc(a = 0.5, mode, intvl)  
    fval[2] = beta_calc(a = 50, mode, intvl)  
  
    z = uniroot(beta_calc, interval=c(0.5, 50), mode , intvl, lower = 0.5, upper=50,  
f.lower=fval[1], f.upper=fval[2])  
    param = vector(mode="numeric", length=2)  
    P1 ∝ Beta(a,b) avec param[1]=a, param[2]=b  
    param[1] = z$root  
    param[2] = ((param[1] - 1)/mode) - (param[1] - 2)  
  
    return(param)  
  }  
}
```

Algorithme du calcul des paramètres de la loi normal :

Fonction pour trouve la valeur de $1/\sigma$ tel que $P(p_2 - p_1 > -c_2) = \text{gamma}$ la réponse présentée par 1 - Q4.

On utilise l'intégrale de Simpson pour intégrer sur la loi jointe de (p_2, p_1) .

Function inputs : $\text{sigmainv} = 1/\sigma$ où σ est l'écart-type *a priori* de θ

$\text{mu_sigma} = \text{mu}/\text{sigma}$, ratio de la moyenne et l'écart-type *a priori* de θ

a, b = paramètres de la loi *a priori* de P_1

gamma = valeur de $P(p_E - p_C > -c_2)$ élicitée par les experts (=1-Q4)

c_2 = marge de non-infériorité de l'essai

Function returns : $P(p_E - p_C > -c_2; 1/\sigma) - \text{gamma}$.

```
calc_thetavar <- function(sigmainv, mu_sigma, a, b, gamma, c2) {
```

Set up grid of equally spaced points over $[0,1]$ for p_1 and over $[\text{max}0, p_1 - c_2, 1]$ for p_2

```
gridc = seq(0.001, 0.999, by=0.001)
```

```
midp1 = (0.00001 + 0.001)*0.5
```

```
midp2 = (0.99999 + 0.999)*0.5
```

```
gridc = append(c(0.00001, midp1), gridc)
```

```
gridc = append(gridc, c(midp2, 0.99999))
```

```
lc = length(gridc)
```

```
wc = vector(mode="numeric", length=lc)
```

```
wc[1] = (gridc[3]-gridc[1])/6.0
```

```
wc[lc] = (gridc[lc] - gridc[lc-2])/6.0
```

```
for(i in seq(2, lc-1, by=2))
```

```
wc[i] = 4*(gridc[i+1] - gridc[i-1])/6.0
```

```
for(i in seq(3, lc-2, by=2))
```

```
wc[i] = (gridc[i+2] - gridc[i-2])/6.0
```

```
int = vector(mode = "numeric", length = lc)
```

```
int1 = as.double(0)
```

```
for(i in 1 :lc){
for each mesh point for pc, integrate joint density over the interval [max0, p1 - c2, 1]
upp = max(0, gridc[i] - c2)
if(upp <= 0){
int[i] = dbeta(gridc[i], shape1=a, shape2=b, ncp=0, log=FALSE)

else{

create a mesh for p2 over the interval [upp, 1]

m = floor((0.999 - upp)/0.002)
u1 = upp + m*0.002
midp1 = (u1 + 0.99999)/2.0
gride = as.double(append(seq(upp, u1, by = 0.001), append(c(midp1), 0.99999) ))
le = length(gride)
if(floor(le/2.0) == (le/2.0))
stop("mesh for MTX remission rate contains even number of elements when odd number
are expected.")

we = vector(mode="numeric", length=le)
we[1] = (gride[3]-gride[1])/6.0
we[le] = (gride[le] - gride[le-2])/6.0
for(j in seq(2, le-1, by=2))
we[j] = 4*(gride[j+1] - gride[j-1])/6.0

for(j in seq(3, (le-2), by=2))
we[j] = (gride[j+2] - gride[j-2])/6.0
```

Evaluating joint density for (p2, p1) at vector of values of p2 for each given gridc[i]

```

dens = vector(mode="numeric", length=le)
dens1 = vector(mode="numeric", length=le)
dens = (gridc[i](a - 1)) * ((1 - gridc[i])(b - 1))/(gride * (1 - gride))
dens1 = -0.5 * ((sigmainv * log(gride * (1 - gridc[i])/(gridc[i] * (1 - gride))) -
mu_sigma)2)
dens = dens * exp(dens1)
int[i] = sum(we * dens)

int1 = sum(wc*int)*sigmainv/(beta(a,b)*sqrt(2*pi))
return(int1 - gamma)

```

Function to identify parameters of a prior normal distribution for theta $N(\mu, \sigma^2)$
 Function inputs : $\pi_1 = P(pE > pC)$ as identified by Day 1 Elicitation Q3
 $\gamma = P(pE - pC > -c_2)$ as identified by 1 - Q4
 a, b = parameters of prior distribution for p_1 c_2 = non-inferiority margin for the trial.
 Function returns : vector with $\text{param}[1] = \mu$ (prior mean of theta), $\text{param}[2] = \sigma^2$ (prior variance of theta).

```
prior_theta <- function(pi1, gamma, a, b, c2)
mu_sigma = qnorm(pi1, mean = 0, sd = 1, lower.tail = TRUE)
fval = vector(mode = "numeric", length = 2)
l1 = as.double(0.01)
u1 = as.double(5)

fval[1] = calc_thetavar(l1, mu_sigma, a, b, gamma, c2)
fval[2] = calc_thetavar(u1, mu_sigma, a, b, gamma, c2)

if(identical(sign(fval[1]), sign(fval[2])))
cat("Given Beta prior for pC and answers to elicitation questions Q3 and Q4, we cannot
determine a Normal prior distribution for the log-odds ratio. ")
cat("Please revise either the answer to elicitation Q3 or Q4. ")
stop("Given Beta prior for p1 and answers to elicitation questions Q3 and Q4, cannot
determine Normal prior distribution for log-odds ratio")

Search variable here is 1/sigma, which is 1 over the prior sd of theta.
z = uniroot(calc_thetavar, interval=c(l1, u1), mu_sigma, a, b, gamma, c2, lower =
l1, upper = u1, f.lower = fval[1], f.upper = fval[2])
param = vector(mode = "numeric", length = 2)
theta ~ normal(mu, sigma^2), avecparam[1] = mu, param[2] = sigma^2
param[1] = mu_sigma/z$root
param[2] = (1/z$root)^2

return(param)
```


Bibliographie

- [Abrahamyan et al., 2016] Abrahamyan, L., Feldman, B. M., Tomlinson, G., Faughnan, M. E., Johnson, S. R., Diamond, I. R., and Gupta, S. (2016). Alternative designs for clinical trials in rare diseases. *American Journal of Medical Genetics Part C : Seminars in Medical Genetics*, 172(4) :313–331. 3 citations pages [32](#), [34](#), et [43](#)
- [Abrahamyan et al., 2011] Abrahamyan, L., Li, C. S., Beyene, J., Willan, A. R., and Feldman, B. M. (2011). Survival distributions impact the power of randomized placebo-phase design and parallel groups randomized clinical trials. *Journal of Clinical Epidemiology*, 64(3) :286–292. Cité page [34](#)
- [Albert et al., 2012] Albert, I., Donnet, S., Guihenneuc-Jouyau, C., Low-Choy, S., Mengersen, K., and Rousseau, J. (2012). Combining expert opinions in prior elicitation. *Bayesian Analysis*, 7(3) :503–532. Cité page [57](#)
- [Apostolakis, 1981] Apostolakis, G. (1981). Bayesian Methods in Risk Assessment. In *Advances in Nuclear Science and Technology*, Advances in Nuclear Science and Technology, pages 415–465. Springer, Boston, MA. DOI : 10.1007/978-1-4613-9919-3_5. Cité page [53](#)
- [Aronson, 2005] Aronson, J. K. (2005). Biomarkers and surrogate endpoints. *British Journal of Clinical Pharmacology*, 59(5) :491–494. Cité page [44](#)
- [Ayyub, 2000] Ayyub, B. M. (2000). *Methods for Expert-Opinion Elicitation of Probabilities and Consequences for Corps Facilities*. PhD thesis. 2 citations pages [57](#) et [101](#)
- [Berger, 1985a] Berger, J. O. (1985a). Bayesian analysis. In *Statistical Decision Theory and Bayesian Analysis*, Springer Series in Statistics, pages 118–307. Springer New York. DOI : 10.1007/978-1-4757-4286-2_4. 4 citations pages [52](#), [66](#), [68](#), et [135](#)
- [Berger, 1985b] Berger, J. O. (1985b). Prior information and subjective probability. In *Statistical Decision Theory and Bayesian Analysis*, Springer Series in Statistics, pages 74–117. Springer New York. DOI : 10.1007/978-1-4757-4286-2_3. 3 citations pages [52](#), [62](#), et [63](#)

- [Berlin, 2010] Berlin, J. A. (2010). N-of-1 clinical trials should be incorporated into clinical practice. *Journal of Clinical Epidemiology*, 63(12) :1283–1284. Cité page [30](#)
- [Berry, 2006] Berry, D. A. (2006). Bayesian clinical trials. *Nature Reviews. Drug Discovery*, 5(1) :27–36. Cité page [43](#)
- [Billingham et al., 2012] Billingham, L., Malottki, K., and Steven, N. (2012). Small sample sizes in clinical trials : a statistician’s perspective. *Clinical Investigation*, 2(7) :655–657. 2 citations pages [79](#) et [80](#)
- [Brookes et al., 2001] Brookes, S. T., Whitley, E., Peters, T. J., Mulheran, P. A., Egger, M., and Davey Smith, G. (2001). Subgroup analyses in randomised controlled trials : quantifying the risks of false-positives and false-negatives. *Health Technology Assessment (Winchester, England)*, 5(33) :1–56. Cité page [43](#)
- [Brophy et al., 2003] Brophy, J. M., Belisle, P., and Joseph, L. (2003). Evidence for use of coronary stents. a hierarchical bayesian meta-analysis. *Annals of Internal Medicine*, 138(10) :777–786. Cité page [26](#)
- [Brunelli, 2015] Brunelli, M. (2015). Priority vector and consistency. In *Introduction to the Analytic Hierarchy Process*, SpringerBriefs in Operations Research, pages 17–31. Springer International Publishing. DOI : 10.1007/978-3-319-12502-2_2. 6 citations pages [87](#), [88](#), [91](#), [93](#), [94](#), et [97](#)
- [Buckley, 2008] Buckley, B. M. (2008). Clinical trials of orphan medicines. *The Lancet*, 371(9629) :2051–2055. 2 citations pages [51](#) et [77](#)
- [Burton, 1994] Burton, P. R. (1994). Helping doctors to draw appropriate inferences from the analysis of medical studies. *Statistics in Medicine*, 13(17) :1699–1713. Cité page [80](#)
- [Caporali et al., 2004] Caporali, R., Cimmino, M. A., Ferraccioli, G., Gerli, R., Klersy, C., Salvarani, C., Montecucco, C., and Systemic Vasculitis Study Group of the Italian Society for Rheumatology (2004). Prednisone plus methotrexate for polymyalgia rheumatica : a randomized, double-blind, placebo-controlled trial. *Annals of Internal Medicine*, 141(7) :493–500. 4 citations pages [142](#), [143](#), [144](#), et [145](#)
- [Cepeda et al., 2003] Cepeda, M. S., Boston, R., Farrar, J. T., and Strom, B. L. (2003). Comparison of logistic regression versus propensity score when the number of events is low and there are multiple confounders. *American journal of epidemiology*, 158(3) :280–287. Cité page [45](#)
- [Chaloner and Rhame, 2001] Chaloner, K. and Rhame, F. S. (2001). Quantifying and documenting prior beliefs in clinical trials. *Statistics in medicine*, 20(4) :581–600. Cité page [110](#)

- [Choy et al., 2009] Choy, S. L., O’Leary, R., and Mengersen, K. (2009). Elicitation by design in ecology : using expert opinion to inform priors for bayesian statistical models. *Ecology*, 90(1) :265–277. Cité page [99](#)
- [Clarke, 2008] Clarke, C. E. (2008). Are delayed-start design trials to show neuroprotection in Parkinson’s disease fundamentally flawed? *Movement Disorders*, 23(6) :784–789. Cité page [24](#)
- [Cole et al., 2011] Cole, J. A., Taylor, J. S., Hangartner, T. N., Weinreb, N. J., Mistry, P. K., and Khan, A. (2011). Reducing selection bias in case-control studies from rare disease registries. *Orphanet Journal of Rare Diseases*, 6 :61. Cité page [44](#)
- [Cooke, 1991] Cooke, R. M. (1991). *Experts in Uncertainty : Opinion and Subjective Probability in Science*. Oxford University Press. 2 citations pages [102](#) et [103](#)
- [Cornu et al., 2013] Cornu, C., Kassai, B., Fisch, R., Chiron, C., Alberti, C., Guerrini, R., Rosati, A., Pons, G., Tiddens, H., Chabaud, S., Caudri, D., Ballot, C., Kurbatova, P., Castellan, A.-C., Bajard, A., and Nony, P. (2013). Experimental designs for small randomised clinical trials : an algorithm for choice. *Orphanet Journal of Rare Diseases*, 8 :48. Cité page [31](#)
- [Crawford, 1987] Crawford, G. B. (1987). The geometric mean procedure for estimating the scale of a judgement matrix. *Mathematical Modelling*, 9(3) :327–334. Cité page [88](#)
- [D’Agostino, 2009] D’Agostino, R. B. S. (2009). The delayed start study design. *New England Journal of Medicine*, 361(13) :1304–1306. 2 citations pages [25](#) et [26](#)
- [David J. Spiegelhalter, 2003] David J. Spiegelhalter, Keith R. Abrams, J. P. M. (2003). Bayesian approaches to clinical trials and health-care evaluation. 3 citations pages [78](#), [80](#), et [81](#)
- [Edwards et al., 1997] Edwards, S. J., Lilford, R. J., Braunholtz, D., and Jackson, J. (1997). Why "underpowered" trials are not necessarily unethical. *Lancet (London, England)*, 350(9080) :804–807. Cité page [45](#)
- [Efficiency, 1987] Efficiency (1987). Efficiency of plasma exchange in Guillain-Barré syndrome : role of replacement fluids. French Cooperative Group on Plasma Exchange in Guillain-Barré syndrome. *Annals of Neurology*, 22(6) :753–761. Cité page [73](#)
- [Efron, 1971] Efron, B. (1971). Forcing a sequential experiment to be balanced. *Biometrika*, 58(3) :403–417. Cité page [39](#)
- [Epstein and Sladky, 1990] Epstein, M. A. and Sladky, J. T. (1990). The role of plasmapheresis in childhood Guillain-Barré syndrome. *Annals of Neurology*, 28(1) :65–69. Cité page [74](#)

- [Feldman et al., 2001] Feldman, B., Wang, E., Willan, A., and Szalai, J. P. (2001). The randomized placebo-phase design for clinical trials. *Journal of Clinical Epidemiology*, 54(6) :550–557. Cité page [33](#)
- [Fleming and Powers, 2012] Fleming, T. R. and Powers, J. H. (2012). Biomarkers and surrogate endpoints in clinical trials. *Statistics in Medicine*, 31(25) :2973–2984. Cité page [44](#)
- [Freeman, 1989] Freeman, P. R. (1989). The performance of the two-stage analysis of two-treatment, two-period crossover trials. *Statistics in Medicine*, 8(12) :1421–1432. Cité page [27](#)
- [Fumika, 2004] Fumika, O. (2004). *A literature review on the use of expert opinion in probabilistic risk analysis*. World Bank Washington, DC. 2 citations pages [102](#) et [103](#)
- [Gagne et al., 2014] Gagne, J. J., Thompson, L., O’Keefe, K., and Kesselheim, A. S. (2014). Innovative research methods for studying treatments for rare diseases : methodological review. *BMJ*, 349 :g6802. Cité page [45](#)
- [Gallin et al., 2003] Gallin, J. I., Alling, D. W., Malech, H. L., Wesley, R., Koziol, D., Marciano, B., Eisenstein, E. M., Turner, M. L., DeCarlo, E. S., Starling, J. M., and Holland, S. M. (2003). Itraconazole to prevent fungal infections in chronic granulomatous disease. *New England Journal of Medicine*, 348(24) :2416–2422. Cité page [45](#)
- [Goodman, 1994] Goodman, S. N. (1994). The use of predicted confidence intervals when planning experiments and the misuse of power when interpreting results. *Annals of Internal Medicine*, 121(3) :200. Cité page [45](#)
- [Goodman and Sladky, 2005] Goodman, S. N. and Sladky, J. T. (2005). A bayesian approach to randomized controlled trials in children utilizing information from adults : the case of Guillain-Barre. *Clinical Trials*, 2(4) :305–310. 3 citations pages [47](#), [72](#), et [81](#)
- [Graf et al., 1999] Graf, W. D., Katz, J. S., Eder, D. N., Smith, A. J., and Chun, M. R. (1999). Outcome in severe pediatric Guillain-Barré syndrome after immunotherapy or supportive care. *Neurology*, 52(7) :1494–1497. Cité page [74](#)
- [Greenhouse and Seltman, 2005] Greenhouse, J. B. and Seltman, H. (2005). Using prior distributions to synthesize historical evidence : comments on the Goodman-Sladky case study IVIg in Guillain-Barré syndrome. *Clinical Trials*, 2(4) :311–318. Cité page [47](#)
- [Griggs et al., 2009] Griggs, R. C., Batshaw, M., Dunkle, M., Gopal-Srivastava, R., Kaye, E., Krischer, J., Nguyen, T., Paulus, K., and Merkel, P. A. (2009). Clinical research for rare disease : Opportunities, challenges, and solutions. *Molecular genetics and metabolism*, 96(1) :20–26. 2 citations pages [26](#) et [31](#)

- [Gsteiger et al., 2013] Gsteiger, S., Neuenschwander, B., Mercier, F., and Schmidli, H. (2013). Using historical control information for the design and analysis of clinical trials with overdispersed count data. *Statistics in Medicine*, 32(21) :3609–3622. Cité page [72](#)
- [Gupta et al., 2011] Gupta, S., Faughnan, M. E., Tomlinson, G. A., and Bayoumi, A. M. (2011). A framework for applying unfamiliar trial designs in studies of rare diseases. *Journal of Clinical Epidemiology*, 64(10) :1085–1094. 2 citations pages [26](#) et [27](#)
- [Halpern SD et al., 2002] Halpern SD, Karlawish JT, and Berlin JA (2002). The continuing unethical conduct of underpowered clinical trials. *JAMA*, 288(3) :358–362. Cité page [46](#)
- [Hampson et al., 2014] Hampson, L. V., Whitehead, J., Eleftheriou, D., and Brogan, P. (2014). Bayesian methods for the design and interpretation of clinical trials in very rare diseases. *Statistics in Medicine*, 33(24) :4186–4201. 5 citations pages [13](#), [72](#), [108](#), [110](#), et [115](#)
- [Hoffman et al., 2002] Hoffman, G. S., Cid, M. C., Hellmann, D. B., Guillevin, L., Stone, J. H., Schousboe, J., Cohen, P., Calabrese, L. H., Dickler, H., Merkel, P. A., Fortin, P., Flynn, J. A., Locker, G. A., Easley, K. A., Schned, E., Hunder, G. G., Sneller, M. C., Tuggle, C., Swanson, Howard, J., Lopez-Soto, A., Bork, D., Hoffman, D. B., Kalunian, K., Klashman, D., Wilke, W. S., Scheetz, R. J., Mandell, B. F., Fessler, B. J., Kosmorsky, G., Prayson, R., Luqmani, R. A., Nuki, G., McRorie, E., Sherrer, Y., Baca, S., Walsh, B., Ferland, D., Soubrier, M., Choi, H. K., Gross, W., Segal, A. M., Ludivico, C., Puechal, X., and International Network for the Study of Systemic Vasculitides (2002). A multicenter, randomized, double-blind, placebo-controlled trial of adjuvant methotrexate treatment for giant cell arteritis. *Arthritis and Rheumatism*, 46(5) :1309–1318. 4 citations pages [125](#), [126](#), [133](#), et [134](#)
- [Hu and Rosenberger, 2006] Hu, F. and Rosenberger, W. F. (2006). *The Theory of Response-Adaptive Randomization in Clinical Trials*. John Wiley & Sons. Google-Books-ID : CmkFQ2Qibk8C. 2 citations pages [36](#) et [37](#)
- [Hughes, 1993] Hughes, M. D. (1993). Reporting Bayesian analyses of clinical trials. *Statistics in Medicine*, 12(18) :1651–1663. Cité page [80](#)
- [Ishizaka and Labib, 2009] Ishizaka, A. and Labib, A. (2009). Analytic hierarchy process and expert choice : Benefits and limitations. *OR Insight*, 22(4) :201–220. Cité page [97](#)
- [Jansen et al., 1993] Jansen, P. W., Perkin, R. M., and Ashwal, S. (1993). Guillain-Barré syndrome in childhood : natural course and efficacy of plasmapheresis. *Pediatric Neurology*, 9(1) :16–20. Cité page [74](#)

- [Johnson et al., 2011] Johnson, S. R., Granton, J. T., Tomlinson, G. A., Grosbein, H. A., Hawker, G. A., and Feldman, B. M. (2011). Effect of warfarin on survival in scleroderma-associated pulmonary arterial hypertension (ssc-pah) and idiopathic pah. belief elicitation for bayesian priors. *The Journal of rheumatology*, 38(3) :462–469. Cité page [110](#)
- [Jover et al., 2001] Jover, J. A., Hernandez-Garcia, C., Morado, I. C., Vargas, A., and Fernandez-Gutierrez, B. (2001). Combined treatment of giant-cell arteritis with methotrexate and prednisone. a randomized, double-blind, placebo-controlled trial. *Annals of Internal Medicine*, 134(2) :106–114. 8 citations pages [125](#), [126](#), [132](#), [134](#), [142](#), [143](#), [144](#), et [145](#)
- [Katz, 2009] Katz, N. (2009). Enriched enrollment randomized withdrawal trial designs of analgesics : focus on methodology. *The Clinical Journal of Pain*, 25(9) :797–807. Cité page [32](#)
- [Khatri et al., 1990] Khatri, B. O., Robert Flamini, J., Baruah, J. K., Dobyns, W. B., and Konkol, R. J. (1990). Plasmapheresis with acute inflammatory polyneuropathy. *Pediatric Neurology*, 6(1) :17–19. Cité page [74](#)
- [Knapp, 1996] Knapp, T. R. (1996). The overemphasis on power analysis. *Nursing Research*, 45(6) :379–381. Cité page [45](#)
- [Koczkodaj, 1993] Koczkodaj, W. W. (1993). A new definition of consistency of pairwise comparisons. *Mathematical and Computer Modelling*, 18(7) :79–84. Cité page [94](#)
- [Korinthenberg and Mönting, 1996] Korinthenberg, R. and Mönting, J. S. (1996). Natural history and treatment effects in Guillain-Barré syndrome : a multicentre study. *Archives of Disease in Childhood*, 74(4) :281–287. Cité page [74](#)
- [Lagakos, 2003] Lagakos, S. W. (2003). Clinical trials and rare diseases. *The New England Journal of Medicine*, 348(24) :2455–2456. Cité page [26](#)
- [Lamont et al., 1991] Lamont, P. J., Johnston, H. M., and Berdoukas, V. A. (1991). Plasmapheresis in children with Guillain-Barré syndrome. *Neurology*, 41(12) :1928–1931. Cité page [74](#)
- [Lilford et al., 1995] Lilford, R. J., Thornton, J., and Braunholtz, D. (1995). Clinical trials and rare diseases : a way out of a conundrum. *Bmj*, 311(7020) :1621–1625. 4 citations pages [47](#), [72](#), [78](#), et [79](#)
- [Lillie et al., 2011] Lillie, E. O., Patay, B., Diamant, J., Issell, B., Topol, E. J., and Schork, N. J. (2011). The n-of-1 clinical trial : the ultimate strategy for individualizing medicine? *Personalized medicine*, 8(2) :161–173. Cité page [30](#)

- [Liu-Seifert et al., 2015] Liu-Seifert, H., Siemers, E., Holdridge, K. C., Andersen, S. W., Lipkovich, I., Carlson, C., Sethuraman, G., Hoog, S., Hayduk, R., Doody, R., and Aisen, P. (2015). Delayed-start analysis : Mild alzheimer's disease patients in solanezumab trials, 3.5 years. *Alzheimer's & Dementia : Translational Research & Clinical Interventions*, 1(2) :111–121. 2 citations pages [24](#) et [25](#)
- [Maclure, 1991] Maclure, M. (1991). The case-crossover design : a method for studying transient effects on the risk of acute events. *American Journal of Epidemiology*, 133(2) :144–153. 2 citations pages [26](#) et [27](#)
- [Mahajan and Gupta, 2010] Mahajan, R. and Gupta, K. (2010). Adaptive design clinical trials : Methodology, challenges and prospect. *Indian Journal of Pharmacology*, 42(4) :201–207. Cité page [35](#)
- [Mann, 2003] Mann, C. J. (2003). Observational research methods. research design ii : cohort, cross sectional, and case-control studies. *Emergency Medicine Journal*, 20(1) :54–60. Cité page [26](#)
- [M.D, 1994] M.D, R. J. T. (1994). Special study designs : early escape, enrichment, studies in non-responders. *Communications in Statistics - Theory and Methods*, 23(2) :499–531. Cité page [33](#)
- [Morita et al., 2008] Morita, S., Thall, P. F., and Müller, P. (2008). Determining the Effective Sample Size of a Parametric Prior. *Biometrics*, 64(2) :595–602. Cité page [117](#)
- [Nason and Follmann, 2010] Nason, M. and Follmann, D. (2010). Design and analysis of crossover trials for absorbing binary endpoints. *Biometrics*, 66(3) :958–965. Cité page [27](#)
- [Neuenschwander et al., 2010] Neuenschwander, B., Capkun-Niggli, G., Branson, M., and Spiegelhalter, D. J. (2010). Summarizing historical information on controls in clinical trials. *Clinical Trials*, 7(1) :5–18. 2 citations pages [72](#) et [118](#)
- [Nony et al., 2014] Nony, P., Kurbatova, P., Bajard, A., Malik, S., Castellan, C., Chabaud, S., Volpert, V., Eymard, N., Kassai, B., and Cornu, C. (2014). A methodological framework for drug development in rare diseases. *Orphanet Journal of Rare Diseases*, 9. Cité page [31](#)
- [O'Hagan et al., 2006] O'Hagan, A., Buck, C. E., Daneshkhah, A., Eiser, J. R., Garthwaite, P. H., Jenkinson, D. J., Oakley, J. E., and Rakow, T. (2006). Wiley : Uncertain judgements : Eliciting experts probabilities. 4 citations pages [57](#), [102](#), [103](#), et [105](#)
- [Olanow et al., 2009] Olanow, C. W., Rascol, O., Hauser, R., Feigin, P. D., Jankovic, J., Lang, A., Langston, W., Melamed, E., Poewe, W., Stocchi, F., et al. (2009). A double-

- blind, delayed-start trial of rasagiline in parkinson's disease. *New England Journal of Medicine*, 361(13) :1268–1278. Cité page [26](#)
- [Paradiso et al., 1999] Paradiso, G., Tripoli, J., Galicchio, S., and Fejerman, N. (1999). Epidemiological, clinical, and electrodiagnostic findings in childhood Guillain-Barré syndrome : a reappraisal. *Annals of Neurology*, 46(5) :701–707. Cité page [74](#)
- [Parkinson, 2004] Parkinson (2004). A controlled, randomized, delayed-start study of rasagiline in early Parkinson disease. *Archives of Neurology*, 61(4) :561–566. Cité page [26](#)
- [Plasma, 1997] Plasma (1997). Randomised trial of plasma exchange, intravenous immunoglobulin, and combined treatments in Guillain-Barré syndrome. Plasma Exchange/Sandoglobulin Guillain-Barré Syndrome Trial Group. *Lancet (London, England)*, 349(9047) :225–230. Cité page [73](#)
- [Plasmapheresis, 1985] Plasmapheresis (1985). Plasmapheresis and acute Guillain-Barré syndrome. The Guillain-Barré syndrome Study Group. *Neurology*, 35(8) :1096–1104. Cité page [73](#)
- [Pocock and Simon, 1975] Pocock, S. J. and Simon, R. (1975). Sequential treatment assignment with balancing for prognostic factors in the controlled clinical trial. *Biometrics*, 31(1) :103–115. Cité page [40](#)
- [Pöyhönen et al., 1997] Pöyhönen, M. A.yhönen, M. A., Hämälä inen, R. P., and Salo, A. A. (1997). An experiment on the numerical modelling of verbal ratio statements. *Journal of Multi-Criteria Decision Analysis*, 6(1) :1–10. Cité page [95](#)
- [Rahme et al., 2000] Rahme, E., Joseph, L., and Gyorkos, T. W. (2000). Bayesian sample size determination for estimating binomial parameters from data subject to misclassification. *Journal of the Royal Statistical Society : Series C (Applied Statistics)*, 49(1) :119–128. Cité page [85](#)
- [Robert, 2007] Robert, C. P. (2007). *The Bayesian Choice*. Springer Texts in Statistics. Springer New York, New York, NY. 3 citations pages [52](#), [62](#), et [66](#)
- [Rosenbaum and Rubin, 1983] Rosenbaum, P. R. and Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1) :41–55. Cité page [45](#)
- [Rosenberger and Lachin, 2016] Rosenberger, W. F. and Lachin, J. M. (2016). Randomization in clinical trials : Theory and practice. Cité page [38](#)
- [Saaty, 1987] Saaty, R. W. (1987). The analytic hierarchy process - what it is and how it is used. *Mathematical Modelling*, 9(3) :161–176. 3 citations pages [94](#), [95](#), et [98](#)

- [Saaty, 1977] Saaty, T. L. (1977). A scaling method for priorities in hierarchical structures. *Journal of Mathematical Psychology*, 15(3) :234–281. 3 citations pages [86](#), [88](#), et [93](#)
- [Saaty and Hu, 1998] Saaty, T. L. and Hu, G. (1998). Ranking by eigenvector versus other methods in the analytic hierarchy process. *Applied Mathematics Letters*, 11(4) :121–125. Cité page [87](#)
- [Saaty and Vargas, 1984] Saaty, T. L. and Vargas, L. G. (1984). Comparison of eigenvalue, logarithmic least squares and least squares methods in estimating ratios. *Mathematical Modelling*, 5(5) :309–324. 2 citations pages [87](#) et [93](#)
- [Samaniego, 2010] Samaniego, F. J. (2010). *A Comparison of the Bayesian and Frequentist Approaches to Estimation*. Springer Science & Business Media. Google-Books-ID : WGaKMfKm0sAC. Cité page [53](#)
- [Sándor, 2008] Sándor, B. (2008). Solution of the least squares method problem of pairwise comparison matrices. *Central European Journal of Operations Research*, 16(4) :345. 2 citations pages [89](#) et [91](#)
- [Schmidli et al., 2013] Schmidli, H., Wandel, S., and Neuenschwander, B. (2013). The network meta-analytic-predictive approach to non-inferiority trials. *Statistical Methods in Medical Research*, 22(2) :219–240. Cité page [72](#)
- [Senn, 2006] Senn, S. (2006). Cross-over trials in statistics in medicine : the first 25 years. *Statistics in Medicine*, 25(20) :3430–3442. Cité page [27](#)
- [Shook, 2010] Shook, S. L. (2010). *The randomized placebo-phase design : Evaluation, interim monitoring and analysis*. PhD thesis, University of Pittsburgh. Cité page [34](#)
- [Simes, 1995] Simes, R. J. (1995). Prospective meta-analysis of cholesterol-lowering studies : the prospective pravastatin pooling (ppp) project and the cholesterol treatment trialists (ctt) collaboration. *The American Journal of Cardiology*, 76(9) :122C–126C. Cité page [45](#)
- [Spiera et al., 2001] Spiera, R. F., Mitnick, H. J., Kupersmith, M., Richmond, M., Spiera, H., Peterson, M. G., and Paget, S. A. (2001). A prospective, double-blind, randomized, placebo controlled trial of methotrexate in the treatment of giant cell arteritis (GCA). *Clinical and Experimental Rheumatology*, 19(5) :495–501. 4 citations pages [125](#), [126](#), [133](#), et [134](#)
- [Straube et al., 2008] Straube, S., Derry, S., McQuay, H. J., and Moore, R. A. (2008). Enriched enrolment : definition and effects of enrichment and dose in trials of pregabalin and gabapentin in neuropathic pain. A systematic review. *British Journal of Clinical Pharmacology*, 66(2) :266–275. Cité page [32](#)

- [Sun and Garrison, 2012] Sun, P. and Garrison, L. P. (2012). Retrospective outcomes studies for orphan diseases : challenges and opportunities. *Current Medical Research & Opinion*, 28(4) :665–667. Cité page [44](#)
- [Taji and Matsumoto, 2006] Taji, K. and Matsumoto, K. (2006). Inverse sensitive analysis of pairwise comparison matrices. *Journal of the Operations Research Society of Japan*, 49(4) :332–341. Cité page [94](#)
- [Tan et al., 2003a] Tan, Dear, K. B. G., Bruzzi, P., and Machin., D. (2003a). Strategy for randomised clinical trials in rare cancers. *BMJ*, 327(7405) :47–49. 3 citations pages [47](#), [79](#), et [80](#)
- [Tan et al., 2003b] Tan, S.-B., Dear, K. B. G., Bruzzi, P., and Machin, D. (2003b). Strategy for randomised clinical trials in rare cancers. *BMJ*, 327(7405) :47–49. Cité page [72](#)
- [Tan et al., 2008] Tan, S.-B., Wee, J., Wong, H.-B., and Machin, D. (2008). Can external and subjective information ever be used to reduce the size of randomised controlled trials? *Contemporary Clinical Trials*, 29(2) :211–219. 3 citations pages [79](#), [80](#), et [110](#)
- [Thurstone, 1927] Thurstone, L. L. (1927). A law of comparative judgment. *Psychological Review*, 34(4) :273–286. Cité page [85](#)
- [Van der Meché and Schmitz, 1992] Van der Meché, F. G. and Schmitz, P. I. (1992). A randomized trial comparing intravenous immune globulin and plasma exchange in Guillain-Barré syndrome. Dutch Guillain-Barré Study Group. *The New England Journal of Medicine*, 326(17) :1123–1129. Cité page [73](#)
- [Wei, 1978] Wei, L. J. (1978). The adaptive biased coin design for sequential experiments. *The Annals of Statistics*, 6(1) :92–100. Cité page [39](#)
- [Wei et al., 1986] Wei, L. J., Smythe, R. T., and Smith, R. L. (1986). K- treatment comparisons with restricted randomization rules in clinical trials. *The Annals of Statistics*, 14(1) :265–274. Cité page [39](#)
- [Wellek and Blettner, 2012] Wellek, S. and Blettner, M. (2012). On the proper use of the crossover design in clinical trials. *Deutsches Arzteblatt International*, 109(15) :276–281. 2 citations pages [27](#) et [28](#)
- [Zucker et al., 2010] Zucker, D. R., Ruthazer, R., and Schmid, C. H. (2010). Individual (N-of-1) trials can be combined to give population comparative treatment effect estimates : Methodologic considerations. *Journal of clinical epidemiology*, 63(12) :1312–1323. Cité page [30](#)