

# Séance 4: à propos de la calibration d'un essai clinique d'antipaludiques

J Gaudart, R Giorgi, JC Thalabard, D Thiam, S Whegang

Septembre 2010

## Introduction

Cette séance aborde dans un premier temps le cas d'un critère binaire simple *succès/ échec* puis le cas d'un critère ordinal. Ces deux situations sont abordées, d'une part, dans le cas d'une évaluation finale unique à un temps donné (J28 dans le protocole OMS actuel), d'autre part, dans le cas d'une évaluation à des temps répétés.

## Cas d'un critère binaire

### Analyse à un temps unique

La cas d'un critère binaire renvoie au cas classique de calibration d'un essai clinique sur la base d'un critère quantitatif comme une différence de pourcentage, un OR ou un RR, avec connaissance ou non préalable du taux de succès attendu dans le bras de référence. De nombreuses fonctions sont disponibles sous R dans les packages suivants

- Package epiR *epi.studysize* : Comparaison de moyennes, proportions, survies
- Package epicalc :
  - *n.for.2p* : Comparaison 2 proportions
  - *n.for.cluster.2p* : Comparaison 2 proportions en cluster
  - *n.for.equi.2p*:
  - *n.for.lqas*
  - *n.for.noninferior.2p*
  - *power.for.2p* : puissance dans le cas 2 proportions
- package Hmisc *samplesize.bin*: calibration pour 2 proportions
- package Design
- Package pwr
  - *ES.h* Taille d'effet pour des proportions
  - *pwr.2p.test* Calcul de puissance pour 2 proportions (même taille)
  - *pwr.2p2n.test* Calcul de puissance pour 2 proportions (tailles différentes)
  - *pwr.anova.test* Calcul de puissance pour des tests dans une ANOVA
  - *pwr.chisq.test* Calcul de puissance pour un test de chi-2
  - *pwr.f2.test* Calcul de puissance dans le cadre d'un glm model
  - *pwr.p.test* Calcul de puissance pour un test sur une proportion (un  $\tilde{A}$ lchantillon)
- package gsDesign Package dévolu aux plans séquentiels mais contenant des fonctions correspondant à des analyses à un temps fini, avec notamment des calibrations pour des situations de non infériorité, qui peuvent être particulièrement intéressantes dans le cadre de l'évaluation des antipaludiques.
  - *nBinomial*: calibration en supériorité ou non- infériorité

- *testBinomial*, *ciBinomial*, *simBinomial* sont des fonctions permettant de calculer la grandeur test pour deux proportions, l'intervalle de confiance de la différence, simuler, respectivement, dans les situations soit de supériorité, soit de non-infériorité.

Nous illustrons cette approche simple à partir des résultats codés en binaire RCPA *versus* non- RCPA des différents essais dont le critère de jugement était à J28.

Essai	Groupe <sub>1</sub> <i>p</i> <sub>1</sub>	Groupe <sub>2</sub> <i>p</i> <sub>2</sub>	Taille totale N calculée
1	43/60	50/61	411
2	55/67	61/69	772
3	52/62	58/61	172
4	53/83	73/85	92
5	73/92	84/91	170

Table 1: Essais J28: Effectif total nécessaire pour mettre en évidence une différence telle qu'observée entre deux groupes de traitement avec un risque  $\alpha = 0.05$  et une puissance de 80%

## Analyse à des temps répétés

Nous présentons ici les grandes lignes de la méthode proposée par Rochon (2). Celle-ci rentre dans le cadre général proposé par Liang & Zeger (6) des équations d'estimations généralisées (GEE), qui ne font pas d'hypothèses particulières sur les distributions sous-jacentes et sont relativement robustes aux mauvaises spécifications sur les moments d'ordre supérieur aux deux premiers moments (espérance et variance).

## Notations et généralités

Dans tout ce qui suit, nous nous placerons dans la situation des essais Cameroun J28, avec comparaison de deux traitements ( $m = 1, 2$ ), sur un critère catégoriel ordinal à 4 classes ( $k = 1, 2, 3, 4$ ), observé à 3 reprises ( $t = 1, 2, 3$ )

Nous désignons par  $Y_{mitk}$  la réponse du sujet  $i$  du groupe de traitement  $m$  au temps  $t$  correspondant à la classe de réponse ordinaire  $k$ , ( $k = 1, \dots, r = 4$ ) et par  $\pi_{mitk}$  la probabilité correspondante qui est telle que  $E(Y_{mitk}) = \pi_{mitk}$ .

Si nous supposons qu'il existe une fonction de lien  $h(\pi_{mitk}) = X_m^T \beta$  où  $X$  correspond aux  $p$  covariables observées, et que la variance peut s'exprimer sous la forme d'une fonction  $V_{mi} = g(\mu_{mi})$ , Liang & Zeger ont montré qu'il était possible d'obtenir des estimateurs consistants des paramètres inconnus du modèle en résolvant l'équation suivante

$$\sum_{\text{sujet}=1}^{\text{sujet}=N} D_{\text{sujet}}^t V_{\text{sujet}}^{-1} (Y_{\text{sujet}} - \mu_{\text{sujet}}) = 0$$

où  $N$  désigne le nombre total de sujets étudiés,  $D_{\text{sujet}} = \frac{\partial \mu_{\text{sujet}}}{\partial \beta} = \frac{\partial \mu_{\text{sujet}}}{\partial h} X_{\text{sujet}} = \Delta_{\text{sujet}} X_{\text{sujet}}$  est une matrice  $T \times p$

$V_{\text{sujet}}$  est une matrice de variance-covariance  $T \times T$

Les données répétées issues d'un même sujet présentent une structure de covariance  $V_m$  qui peut s'écrire

$$V_m = \Psi A_m^{1/2} \cdot R \cdot A_m^{1/2} \quad A_m = \text{diag} [g(\mu_{mt_1}), g(\mu_{mt_2}), g(\mu_{mt_3})]$$

$R$  est une matrice dite *de travail* qui reflète le degré de corrélation entre les mesures chez un même sujet dans le temps

$\Psi$  est un paramètre positif de dispersion relative, qui peut être mis à 1 au stade planification d'étude, sauf informations contraires

### Equations finales

En supposant que tous les sujets appartenant au même groupe de traitement et avec les mêmes covariables ont les mêmes caractéristiques, il est possible de montrer qu'un estimateur du vecteur inconnu des paramètres  $\beta$  est le suivant

$$\hat{\beta} = \left[ \sum_m X_m' W_m X_m \right]^{-1} \left[ \sum_m X_m' W_m h(\mu_m) \right] \quad W_m = \Delta_m \cdot V_m^{-1} \cdot \Delta_m$$

Dans le cas d'un modèle théorique sous-jacent,  $\hat{\beta}$  a pour matrice de covariance

$$\text{cov}(\hat{\beta}) = n^{-1} \cdot \left( \sum_m D_m' \cdot V_m \cdot D_m \right) = n^{-1} \cdot \Omega_M$$

En absence d'un tel modèle sous-jacent théorique, il est possible de proposer une matrice de covariance robuste, dite *sandwich*, pour  $\hat{\beta}$

$$\text{cov}(\hat{\beta}) = n^{-1} \cdot \left( \sum_m \hat{D}_m' \cdot \hat{V}_m \cdot \hat{D}_m \right)^{-1} \left[ \sum_m \hat{D}_m' \cdot \hat{V}_m^{-1} \cdot \hat{\Gamma}_m \cdot \hat{V}_m^{-1} \cdot \hat{D}_m \right] \left( \sum_m \hat{D}_m' \cdot \hat{V}_m \cdot \hat{D}_m \right)^{-1} = n^{-1} \cdot \Omega_R$$

où  $\hat{\Gamma}$  est une estimation de la vraie matrice de variance-covariance des observées  $Y_{\text{sujet}} \Gamma$

### Comparaison d'hypothèses

Si nous associons à l'hypothèse  $H_0$  une certaine combinaison ou expression linéaire  $H$  des paramètres telle que  $H\beta = 0$  versus  $H\beta = \delta$  sous  $H_1$ , la statistique de Wald associée à la grandeur statistique  $H\beta$

$$Q_W = n(H\hat{\beta} - 0)' [H \cdot \Omega \cdot H']^{-1} (H\hat{\beta} - 0)$$

suit, sous  $H_0$ , une loi de Chi-2 centrée à  $p$  degrés de liberté, à laquelle on peut associer un seuil  $S_{1-\alpha}$  pour un niveau de risque  $\alpha$  donné. Sous  $H_1$ ,  $Q_W$  suit une loi chi-2 décentrée de paramètre  $\lambda \approx n(H\hat{\beta} - 0)' [H \cdot \Omega \cdot H']^{-1} (H\hat{\beta} - 0)$ .

La puissance associée au risque de type 2  $\gamma$ , sous  $H_1$  peut s'exprimer sous la forme suivante

$$1 - \gamma = \int_{S_{1-\alpha}}^{\infty} f(x, h, \lambda) dx$$

$h$  désigne dans toutes ces équations le nombre de degrés de liberté qui correspond à la dimension de  $H$

## Algorithme

- **Etape 1** Choix du risque  $\alpha = 0.05$  et de la puissance  $1 - \eta = 0.80$  et du nombre de sous-groupes, dont les groupes de traitements, ici choisi égal à 2
- **Etape 2** Choix du critère principal binaire RCPA/ non RCPA . Choix du nombre de visites dans le temps,  $T = 3$  pour les protocoles J28, correspondant aux jours 14, 21 et 28
- **Etape 3** Choix des probabilités de référence au cours du temps et du gain attendu à tester du nouveau traitement. Choix du lien entre la probabilité de succès / échec et les covariables, sous forme d'un lien logit (*proportional odds*). Nous avons vu que d'autres liens sont possibles (probit, clog-log)
- **Etape 4** Choix des matrices de plan d'expérience (design matrix)  $X_m$

$$X_1 = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \end{bmatrix} \quad X_2 = \begin{bmatrix} 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}$$

- **Etape 5** Spécification de la matrice de corrélation de travail R, qui peut être diagonale (indépendance), symétrique, autorégressive ou avec une forme intermédiaire
- **Etape 6** Calcul du  $\hat{\beta}$  estimé et de sa variance
- **Etape 7** Estimation du  $\lambda$  de décentrage de la loi du Chi-2 sous l'hypothèse alternative: calcul de la puissance *a posteriori* ou algorithme itératif d'estimation de la taille par groupe pour atteindre une puissance donnée

## Application pratique

L'algorithme présenté a été implémenté sous R, qui dispose de fonctions de manipulations matricielles adaptées, ainsi que des fonctions codant les principales lois statistiques, dont la loi de Chi-2 décentrée. Nous donnons ci-dessous un exemple d'appel à la procédure, qui fait appel à deux sources utilitaires qui sont données en annexe.

Les valeurs numériques adoptées correspondent à des observations binaires à 3 temps, avec comme probabilité de succès de référence 0.85 et un delta d'amélioration de 0.1. Le coefficient  $\phi$  de corrélation entre deux temps vaut 0.5 avec une puissance d'atténuation avec le temps de  $\theta = 0.5$

```
> repertoire = "/media/JCT31012009/Enseignement/Stafav/Theses/SWhegang/Programmes_R/"
> Util1 = paste(repertoire, "Gee_SampleSizeBinaire_Utilitaires1.r",
+   sep = "", collapse = NULL)
> Util2 = paste(repertoire, "Gee_SampleSizeBinaire_Utilitaires2.r",
+   sep = "", collapse = NULL)
> source(Util1)
> source(Util2)
> T = 3
> mu = 0.8
> Delta = 0.1
> theta = 0.5
> phi = 0.5
> df = 2
```

```
> alpha = 0.05
> puissance = 0.8
> precision = 1e-04
> lambda0 = Calcul_lambda0(T, mu, Delta, theta, phi)
> Taille(puissance, df, lambda0, (1 - alpha), precision)
```

```
lambda0    n Puissance_atteinte
1 0.0595176 162                0.799988
```

Le nombre de sujets par bras de traitement doit donc être au minimum de  $n = 162$  pour atteindre la puissance de 0.80 avec un risque  $\alpha$  de 0.05

## Cas d'un critère ordinal

### Critère ordinal évalué à un temps unique

#### Approche paramétrique via le modèle à "proportional odds" (5)

Dans le cas d'un critère ordinal univarié, J Whitehead (5) a proposé une approche qui permet soit de calibrer *a priori* un essai lors de l'étape de planification, soit de calculer *a posteriori* la puissance d'un test de comparaison effectué sur un critère ordinal. Cette méthode repose sur le modèle de rapports de côtes proportionnels (*proportional odds*). En pratique, cette méthode est implémentée dans le Package *Hmisc* (Frank Harrell) et correspond à la fonction *posamsize*. Le calcul de la puissance *a posteriori* est donné par la fonction *popower*

#### Utilisation des fonctions *posamsize* et *popower*

Traitement	n inclus	n analysés	ETP	ECT	EPT	RCPA
AQ	179	174	2	0	2	170
AQSP	183	173	0	0	1	172
SP	176	172	12	2	4	154
Total	538	519	14	2	7	496

Table 2: Résumé de l'essai 2003

**Application aux essais 2003** Le calcul de la taille suppose de spécifier dans la fonction *posamsize* un vecteur de probabilité marginale, moyenne pondérée des probabilités des deux bras considérés par exemple, l'odds ratio souhaité être mis en évidence, la fraction de sujets alloués au groupe 1 (nous avons supposé des groupes de même taille), un risque  $\alpha = 0.05$  et la puissance souhaitée  $1 - \beta = 0.80$ .

Le calcul d'une puissance *a posteriori* avec la fonction *popower* nécessite de fournir les probabilités observées dans chaque bras, le nombre de sujets par bras, l'odds ratio de l'étude ainsi que le risque  $\alpha$

**Exemple** En prenant comme bras de référence, le vecteur

$$\pi_{\text{marginal}} = (496/519, 7/519, 2/519, 14/519)$$

```
> library(Hmisc)
> p = c(496/519, 7/519, 2/519, 14/519)
> OR = 2
```

**Calcul de la puissance associée à la comparaison entre les bras AQ et AQSP avec les effectifs utilisés avec un  $OR = 2$**

```
> popower(p, odds.ratio = OR, n1 = 174, n2 = 173, alpha = 0.05)
```

Power: 0.263

Efficiency of design compared with continuous response: 0.127

**Calcul du nombre total de sujets nécessaires pour une comparaison permettant de mettre en évidence un  $OR = 3$  par rapport à la répartition de probabilité de référence**

```
> posamsize(p, odds.ratio = OR, fraction = 0.5, power = 0.8, alpha = 0.05)
```

Total sample size: 1542.1

Efficiency of design compared with continuous response: 0.127

Comparaison	OR	RCPA	EPT	ECT	ETP	N	Puissance
SP vs AQ	0.145	0.936	0.0173	0.0005	0.0406	140	0.993
AQSP vs AQ	1.7	0.985	0.0086	0.0	0.0057	7883	0.087

Table 3: Taille et puissance *a posteriori* des comparaisons par rapport à AQ dans l'essai 2003

### Tableau récapitulatif

#### Approche non paramétrique (4)

La méthode précédemment décrite repose sur une approximation normale classique qui fonctionne en général bien lorsque les effectifs attendus sont suffisants pour rendre cette hypothèse valide. Par contre, dans le cas d'effectifs attendus petits, elle peut être en défaut. Plusieurs auteurs ont proposé des méthodes non paramétriques, dont l'inconvénient vient essentiellement des temps de calculs éventuels pour explorer les combinaisons possibles d'un tableau. Wan et al (4) ont proposé un algorithme capable d'améliorer ces temps de calcul, qui doit être relativement facilement implémentable sous R (à venir!)

Indice i bras Traitement	$n_{i1}$	$n_{i2}$	$n_{i3}$	$n_{i4}$	
i=1	62	0	0	1	$n_{1+} = 63$
i=2	53	2	0	6	$n_{2+} = 61$
$n+$	115	2	0	7	$n = n_{1+} + n_{2+} = 124$

**Notations**  $n_1$  désigne le vecteur ligne (62,0,0,1) et  $n_2$  désigne le vecteur ligne (53,2,0,6). On désigne par  $\pi_i = (\pi_{i1}, \pi_{i2}, \pi_{i3}, \pi_{i4}), i = 1, 2$  les probabilités théoriques de répartition entre les différentes catégories  $j = 1, 2, 3, 4$  de réponses au traitement dans chacun des bras  $i = 1, 2$ .

**Principe de calcul: calcul des puissances conditionnelle/ non conditionnelle** Pour des totaux marginaux lignes et colonnes fixés, il est possible de calculer la probabilité d'avoir observé une répartition donnée d'un couple de vecteurs  $(n_1, n_2)$

$$P(n_1/n_+) = \frac{\frac{n_{1+}!}{n_{11}!n_{12}!n_{13}!n_{14}!} \pi_{11}^{n_{11}} \pi_{12}^{n_{12}} \pi_{13}^{n_{13}} \pi_{14}^{n_{14}} \cdot \frac{n_{2+}!}{n_{21}!n_{22}!n_{23}!n_{24}!} \pi_{21}^{n_{21}} \pi_{22}^{n_{22}} \pi_{23}^{n_{23}} \pi_{24}^{n_{24}}}{P(n_+)}$$

Le numérateur  $P(n_+)$  correspond en fait à la somme des probabilités de tous les tableaux qu'il est possible de construire en respectant les contraintes des totaux marginaux constants correspondant au tableau initial, c'est à dire à la somme sur tous les termes identiques au numérateur pour toutes les répartitions compatibles avec les contraintes marginales.

- Sous l'hypothèse nulle  $H_0$  d'absence de différence entre les 2 bras de traitement  $i = 1, 2$ , les vecteurs de probabilités sont égaux  $\pi_1 = \pi_2$ . L'expression ci-dessus se simplifie et devient

$$P(n_1/n_+|H_0) = \frac{C_{n_+1}^{n_{11}} C_{n_+2}^{n_{12}} C_{n_+3}^{n_{13}} C_{n_+4}^{n_{14}}}{C_n^{n_{1+}}}$$

- Pour une statistique de test T et un risque  $\alpha$  donné, la zone critique de rejet R de  $H_0$  doit satisfaire l'équation suivante

$$R = n_1 \text{ tq } n_1 \text{ admissible \& } \sum_{T \geq t} P(n_1/n_+|H_0) \leq \alpha$$

- Etant donné une hypothèse alternative  $H_1$ , le **pouvoir conditionnel** est donné par

$$\sum_{n_1 \in R} P(n_1/n_+|H_1)$$

- Etant donné une hypothèse alternative  $H_1$ , le **pouvoir non- conditionnel** est donné par

$$1 - \beta = \sum_{n_+} \left( \sum_{n_1 \in R} P(n_1/n_+|H_1) \right)$$

**Cas particulier du modèle *proportional odds*** Dans le cas du modèle *proportional odds*, si on désigne par  $Q_i = (Q_{i1}, Q_{i2}, Q_{i3}, Q_{i4})$  avec  $Q_{i1} = \pi_{i1}, Q_{i2} = \pi_{i1} + \pi_{i2}, Q_{i3} = \pi_{i1} + \pi_{i2} + \pi_{i3}, Q_{i4} = \pi_{i1} + \pi_{i2} + \pi_{i3} + \pi_{i4} = 1$ ) les probabilités cumulées, le modèle s'écrit

$$\text{logit}(Q_{ik}) = L_{ik}, k = 1, 2, 3$$

Sous l'hypothèse qu'entre les deux groupes de traitement ( $i = 1, 2$ ), la différence  $\Delta$  reste constante quelle que soit la catégorie de réponse. Sous l'hypothèse d'absence de différence d'effet entre traitement,  $H_0, \Delta = 0$ , permettant de définir une zone de rejet en utilisant la définition précédente de R. Sous l'hypothèse alternative d'un  $\Delta$  fixé lié à un effet traitement, en se donnant le vecteur des  $\pi_2$ , il est possible de calculer le vecteur des  $\pi_1$ , d'où un algorithme de calcul de la puissance non- conditionnelle.

**Principe de l'algorithme**

- Choix de deux valeurs initiales de  $n_A$  et  $n_B$
- Choix de la répartition en proportion entre les deux groupes  $(n_1, n_2)$ , par exemple  $n/2$



- Tirage au hasard d'une répartition marginale  $n_+$  compatible
  - $n = 100$
  - $pi_2 = c(186/191, 1/191, 1/191, 3/191)$
  - $N = matrix(0, nrow = 1, ncol = 4)$
  - $for(i in 1:n) \{i = sample(1:4)[1]; N[1,i] = N[1,i+1]\}$
- Calcul de R pour un  $\alpha$  donné
  - Calcul de la statistique ordinale  $T = \sum_{j=1}^{j=4} j * n_{1j}$  pour toutes les répartition  $n_1$  compatibles avec  $n_+$
  - Histogramme des T et identification de la valeur de t telle que  $\alpha$  est contrôlé
- Calcul de la puissance  $p = 1 - \beta$  non- conditionnelle f(n) pour  $n = n_A, n_B$
- On pose  $g(n) = f(n) - p$
- Comparaison de  $g(n_A)$  et  $g(n_B)$  et utilisation d'un algorithme par dichotomie pour approcher le zéro de la fonction g.

### Formulaires et Arbre de décision pour calibration de données ordinales, temps unique d'évaluation. d'après SJ Walters (3)

#### Formulaires

	Méthode 1	Méthode 2	Méthode 3
Hypothèses	Variable continue Normale	Variable continue Normale	Variable ordinale OR constant & petit Grand Echantillon
Mesure Effet	<b>Diff. moyennes</b>	<b>Médiane</b>	<b>OR</b>
Test	t- test	U Mann- Whithney	PO
Taille effet	$\Delta = \frac{\mu_T - \mu_C}{\sigma}$	$p_{Noetre} = P(Y > X)$	$OR_{Ord} = \frac{\gamma_{iT}}{1 - \gamma_{iT}} / \frac{\gamma_{iC}}{1 - \gamma_{iC}}$
Taille			
Echantillon	$\frac{2[z_{1-\alpha/2} + z_{1-\beta}]^2}{\Delta^2}$	$\frac{[z_{1-\alpha/2} + z_{1-\beta}]^2}{6(p_{Noetr} - 0.5)^2}$	$\frac{6[z_{1-\alpha/2} + z_{1-\beta}]^2 / (\log(OR_{Ord}))^2}{1 - \sum_{i=1}^{i=k} \pi^2}$

#### Arbre de décision

- Si **disponibilité** d'un jeu de données antérieures
  - Si faible nombre de catégories avec comptage  $< 7$ , *Méthode 3*
  - Si fort nombre de catégories avec faibles comptages
    - \* Si Distribution équilibrée, *Méthode 1*
    - \* Si Distribution déséquilibrée, *Méthode 2*
- Si **absence** de données antérieures
  - Utilitacion d'une méthode bootstrap par simulation: *Méthode 4*
    - \* Si distribution symétrique, t-test
    - \* Si distribution asymétrique, test U Mann-Whitney

## Cas d'un critère ordinal évalué de manière répétée dans le temps

Dans le cas d'une réponse catégorielle répétée, Kim et al (1) ont proposé une méthode de calcul de taille d'échantillon qui généralise la méthode proposée par Rochon dans le cas d'une mesure unique (2). A notre connaissance, cette méthode n'est pas actuellement implémentée sous R.

### Notations et généralités

Dans tout ce qui suit, nous nous placerons dans la situation des essais Cameroun J28, avec comparaison de deux traitements ( $m = 1, 2$ ), sur un critère catégoriel ordinal à 4 classes ( $k = 1, 2, 3, 4$ ), observé à 3 reprises ( $t = 1, 2, 3$ )

Nous désignons par  $Y_{mitk}$  la réponse du sujet  $i$  du groupe de traitement  $m$  au temps  $t$  correspondant à la classe de réponse ordinaire  $k$ , ( $k = 1, \dots, K = 4$ ) et par  $\pi_{mitk}$  la probabilité correspondante qui est telle que  $E(Y_{mitk}) = \pi_{mitk}$ .

En reprenant le modèle à rapport de côtes proportionnel (*proportional odds*) déjà utilisé, si  $Q_{mitk}$  désigne la probabilité cumulée  $\sum_{j=1}^{j=k} \pi_{mitj}$ , nous supposons qu'il existe une fonction de lien  $h(Q_{mitk}) = X^T \cdot \beta$  ( $k = 1, 2, 3 = (K - 1)$ ) où  $X$  correspond aux  $p$  covariables observées, et que la matrice de variance-covariance peut s'exprimer sous la forme d'une fonction  $g(\mu_m)$ , il est possible de se ramener à une formulation identique à celle vue dans le cas binaire, avec une dimensionnalité plus importante, puisque, à chaque temps et dans chaque groupe de traitement, il y a  $(K - 1) = 3$  observations, soit un total sur tous les temps d'observations de  $T(K - 1) = 9$  observations par sujet.

Les matrices *design*  $X_m$  sont des matrices  $(T(K - 1) \times ((K - 1) + 1))$  du fait des  $(K - 1)$  paramètres supplémentaires du modèle correspondant aux valeurs de coupures de la variable latente associée du modèle  $\gamma_1 \leq \gamma_2 \leq \gamma_{K-1}$  en plus de l'effet traitement. Les deux types de matrices  $X_1$  et  $X_2$  sont

$$X_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad X_2 = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

Les matrices  $\Delta_1$  et  $\Delta_2$  sont des matrices diagonales  $T(K - 1) \times T(K - 1)$  constituées à partir des  $Q_{mitk}(1 - Q_{mitk})$ .

La matrice de variance-covariance est plus complexe. Elle se compose de  $T \times T = 9$  blocs carrés de taille  $(K - 1) \times (K - 1) = 3 \times 3$

- $T = 3$  blocs diagonaux correspondant à la répétition de la matrice de variance-covariance  $A_m = (a_{mkj})$  entre les réponses  $(Y_{mitk}, Y_{itj})$  pour un sujet donné à un temps donné  $t$ , où  $a_{mkk} = \pi_{mk}(1 - \pi_{mk})$  et  $a_{mkj} = -\pi_{mj}\pi_{mk}$ , comme déjà vu auparavant.
- 6 blocs anti-diagonaux qui reflètent le degré de covariance entre des catégories de réponses à des temps différents. Il est postulé, ici, que cette dépendance est construite autour de deux paramètres  $\phi_1, \phi_2$ , le premier correspondant à la corrélation entre les réponses entre  $Y_{mitk}$  et  $Y_{misk}$  dans la même classe aux deux instants différents  $t, s$ , le second correspondant à la corrélation entre les réponses entre  $Y_{mitk}$  et  $Y_{misj}$  ( $k \neq j$  aux

deux instants différents  $t, s$ . Cette corrélation est supposée atténuée en fonction de la différence de temps  $|t - s|$  selon un mécanisme identique à celui vu dans le cas binaire en  $\phi_l^{|t-s|^\theta}$ ,  $l = 1, 2$

- Quelques précautions supplémentaires doivent être prises dans le calcul de  $V_m$  en introduisant la transposée des matrices  $A_m$

$$V_m = \Psi t(A_m^{1/2}) \cdot R \cdot A_m^{1/2} \text{ où } A_m = \text{diag} [g(\mu_{mt_1}, \mu_{mt_2}, \mu_{mt_3})]$$

Les matrices  $A_m$  ne sont plus diagonales et les matrices  $A_m^{1/2}$  sont obtenues à partir de leur décomposition singulière et leurs valeurs propres.

### Comparaison d'hypothèses

On retrouve l'expression déjà vue dans le cas binaire pour comparer deux hypothèses  $H_0$  versus  $H_1$  avec la statistique de Wald faisant intervenir une loi de Chi-2 décentrée.

Dans le cas du test d'un effet traitement, la combinaison linéaire des 4 paramètres estimés ne concerne en fait que le seul effet traitement, avec un effet supposé identique sur les logit entre classes. La forme linéaire  $H$  prend donc l'allure suivante

$$H = [0 \ 0 \ 0 \ 1] \Rightarrow H\beta = \beta_4 \text{ qui correspond au seul effet traitement}$$

### Algorithme

Nous retrouvons un algorithme en 7 étapes identiques à celui détaillé dans le cas d'un critère binaire avec cependant des modifications importantes aux étapes suivantes

- **Etape 2** Choix du critère principal ordinal et définition de sa taille,  $K = 4$  dans le cas du critère ordinal OMS.  $T = 3$  dans le cas des essais J28
- **Etape 4** Choix des matrices de plan d'expérience  $X_1, X_2$  de dimension  $T(K - 1) \times (K - 1) + 1 = 9 \times 4$
- **Etape 5** Choix de la matrice de corrélation de travail  $R$  de dimension  $T(K - 1) \times T(K - 1) = 9 \times 9$  avec l'aide des deux coefficients  $\phi_1, \phi_2$  et  $\theta$

### Application aux essais J28

#### Choix d'une distribution de référence entre catégories au cours du temps

#### Implémentation sous R

Nous ne reproduisons ici que les ordres principaux d'appel développés. Les sources utilitaires sont donnés en annexe.

```
> repertoire = "/media/JCT31012009/Enseignement/Stafav/Theses/SWhegang/Programmes_R/"
> Util1 = paste(repertoire, "Gee_SampleSizeOrdinal_Utilitaires1.r",
+   sep = "", collapse = NULL)
> Util2 = paste(repertoire, "Gee_SampleSizeOrdinal_Utilitaires2.r",
+   sep = "", collapse = NULL)
> source(Util1)
> source(Util2)
```

Time	RCPA	LPF	LCF	ETP
T1	58	2	1	0
T2	54	0	2	5
T3	50	2	2	7
Proba				
$\pi_{T1}$	0.950	0.030	0.016	0.004
$\pi_{T2}$	0.871	0.004	0.033	0.082
$\pi_{T3}$	0.821	0.033	0.033	0.115
$\pi_{moy}$	0.881	0.023	0.027	0.069

Table 4: Choix des probabilités théoriques dans le groupe de référence, à partir des valeurs du bras AQ dans les essais J28. Les valeurs NA ont été supposées correspondre à des échecs thérapeutiques. Les valeurs nulles, non observées, ont été mises à 1, en augmentant en parallèle les comptages des catégories représentées

```

> T = 3
> K = 4
> theta = 0
> phi1 = 0.5
> phi2 = 0.2
> phi = c(phi1, phi2)
> df = 4
> alpha = 0.05
> puissance = 0.8
> precision = 1e-04
> mu0 = c(0.8, 0.1, 0.05, 0.05)
> OR = 2
> visu(mu0, OR)

      mu1      mu2  Fu1      Fu2
1 0.80 0.88888889 0.80 0.8888889
2 0.10 0.05847953 0.90 0.9473684
3 0.05 0.02699055 0.95 0.9743590
4 0.05 0.02564103 1.00 1.0000000

      mu1      mu2
[1,] 0.80 0.88888889
[2,] 0.10 0.05847953
[3,] 0.05 0.02699055
[4,] 0.05 0.02564103

      Fu1      Fu2
[1,] 0.80 0.8888889
[2,] 0.90 0.9473684
[3,] 0.95 0.9743590

> lambda0 = t(H %*% beta_hat) * Divis * (H %*% beta_hat)
> Taille(puissance, df, lambda0, (1 - alpha), precision)

      lambda0 n_par_groupe Puissance_atteinte
1 0.1183467          101          0.7999713

```

Valeurs Observées Cameroun,  $OR = 2$

```
> mu0 = c(0.88, 0.02315, 0.02726, 0.06682)
> OR = 2
> visu(mu0, OR)
```

	mu1	mu2	Fu1	Fu2
1	0.88000	0.93617021	0.88000	0.9361702
2	0.02315	0.01294047	0.90315	0.9491107
3	0.02726	0.01483998	0.93041	0.9639507
4	0.06682	0.03604934	1.00000	1.0000000

```
> lambda0 = Calcul_lambda0(T, K, mu0, OR, theta, phi)
> Taille(puissance, df, lambda0, (1 - alpha), precision)
```

	lambda0	n_par_groupe	Puissance_atteinte
1	0.08352062	143	0.7999142

Modification de l'effet traitement  $OR = 4$

```
> OR = 4
> lambda0 = Calcul_lambda0(T, K, mu0, OR, theta, phi)
> Taille(puissance, df, lambda0, (1 - alpha), precision)
```

	lambda0	n_par_groupe	Puissance_atteinte
1	0.1116973	107	0.8000891

## Discussion- Conclusion

La séance a donné un aperçu de méthodes adaptées pour déterminer *a priori*, au stade de planification, la taille d'un essai. Ces méthodes nécessitent la connaissance du type de critère de jugement, qui peut être continu, binaire ou ordinal.

Dans le cas plus particulier d'un critère de jugement évalué à des temps répétés, l'intérêt réside dans la possibilité de mettre en évidence un effet traitement qui apparaîtrait au cours du temps. Cependant, dans ce travail, nous nous sommes limités à l'hypothèse d'un effet traitement constant au cours du temps et identique par catégorie de réponse. D'autres hypothèses pourraient être explorées en fonction des propriétés des combinaisons testées concernant leur effet dans le temps.

A l'heure actuelle, les ACTs ont une proportion de succès complet (RCPA) supérieure à 80%, ce qui laisse peu d'information disponible dans des essais de taille limitée pour une évaluation précise des proportions dans les 3 autres classes. De ce fait, l'illustration donnée reste imparfaite car reposant sur des valeurs de référence très imprécises. Par contre, une meilleure précision pourrait être obtenue en rassemblant un plus grand nombre d'études J28.

Actuellement, compte-tenu de la bonne efficacité d'ensemble des ACTs, les choix entre stratégies médicamenteuses, notamment en terme de santé publique, se posent essentiellement en terme de baisse des effets secondaires qu'ils soient immédiats (tolérance chez l'enfant) ou retardés avec l'apparition des résistances.

# Bibliography

- [1] Hae-Young Kim, John M Williamson, and Cynthia M Lyles. Sample-size calculations for studies with correlated ordinal outcomes. *Stat Med*, 24(19):2977–2987, Oct 2005.
- [2] J. Rochon. Application of gee procedures for sample size calculations in repeated measures experiments. *Stat Med*, 17(14):1643–1658, Jul 1998.
- [3] Stephen J Walters. Sample size and power estimation for studies with health related quality of life outcomes: a comparison of four methods using the sf-36. *Health Qual Life Outcomes*, 2:26, May 2004.
- [4] Shu-Mei Wan, Chien-Hua Wu, Ya-Min Tseng, and Ming-Jie Wang. An improved algorithm for sample size determination of ordinal response by two groups. *Communications in Statistics - Simulation and Computation*, 38:2235–2242, 2009.
- [5] J. Whitehead. Sample size calculations for ordered categorical data. *Stat Med*, 12(24):2257–2271, Dec 1993.
- [6] S. L. Zeger and K. Y. Liang. Longitudinal data analysis for discrete and continuous outcomes. *Biometrics*, 42(1):121–130, Mar 1986.