

Modèles Bayésiens hiérarchiques avec variable latente pour des mesures répétées multiples: Application à la maladie d'Alzheimer et au Diabète

Présenté par Jihane AOUNI

Encadrants: M. Loïc DARCHY et M. Pierre COLIN

Rapporteurs: M. François ROUSSET et Mme Nadine HILGERT

Pour obtenir le Master 2 Biostatistique

Etablissement: Université Montpellier-Faculté des sciences

Lieu du stage: Sanofi-Aventis, Chilly Mazarin



SOMMAIRE

● INTRODUCTION	p. 2
● PROBLEMATIQUE	p. 3
● MODELISATION DE L'ALZHEIMER ET RESULTATS	p. 4
● MODELISATION DU DIABETE ET RESULTATS	p. 12
● CONCLUSION	p. 18
● BIBLIOGRAPHIE	p. 19
● ANNEXES	p. 22

INTRODUCTION

INTRODUCTION

- Alzheimer et Diabète: maladies multidimensionnelles et progressives au cours du temps

INTRODUCTION

- Alzheimer et Diabète: maladies multidimensionnelles et progressives au cours du temps
- Alzheimer: Perte progressive de la mémoire et des fonctions cognitives, altération des paramètres IRM (exemple: diminution du volume de l'hippocampe)

INTRODUCTION

- Alzheimer et Diabète: maladies multidimensionnelles et progressives au cours du temps
- Alzheimer: Perte progressive de la mémoire et des fonctions cognitives, altération des paramètres IRM (exemple: diminution du volume de l'hippocampe)
- Diabète: Titration régulière de l'insuline basée sur le glucose à jeun, contrôle de l'HbA1c (ou hémoglobine glyquée) afin de juger l'équilibre de la glycémie au cours des 2 à 3 mois qui précèdent un dosage sanguin, et répétition des épisodes d'hypoglycémie

INTRODUCTION

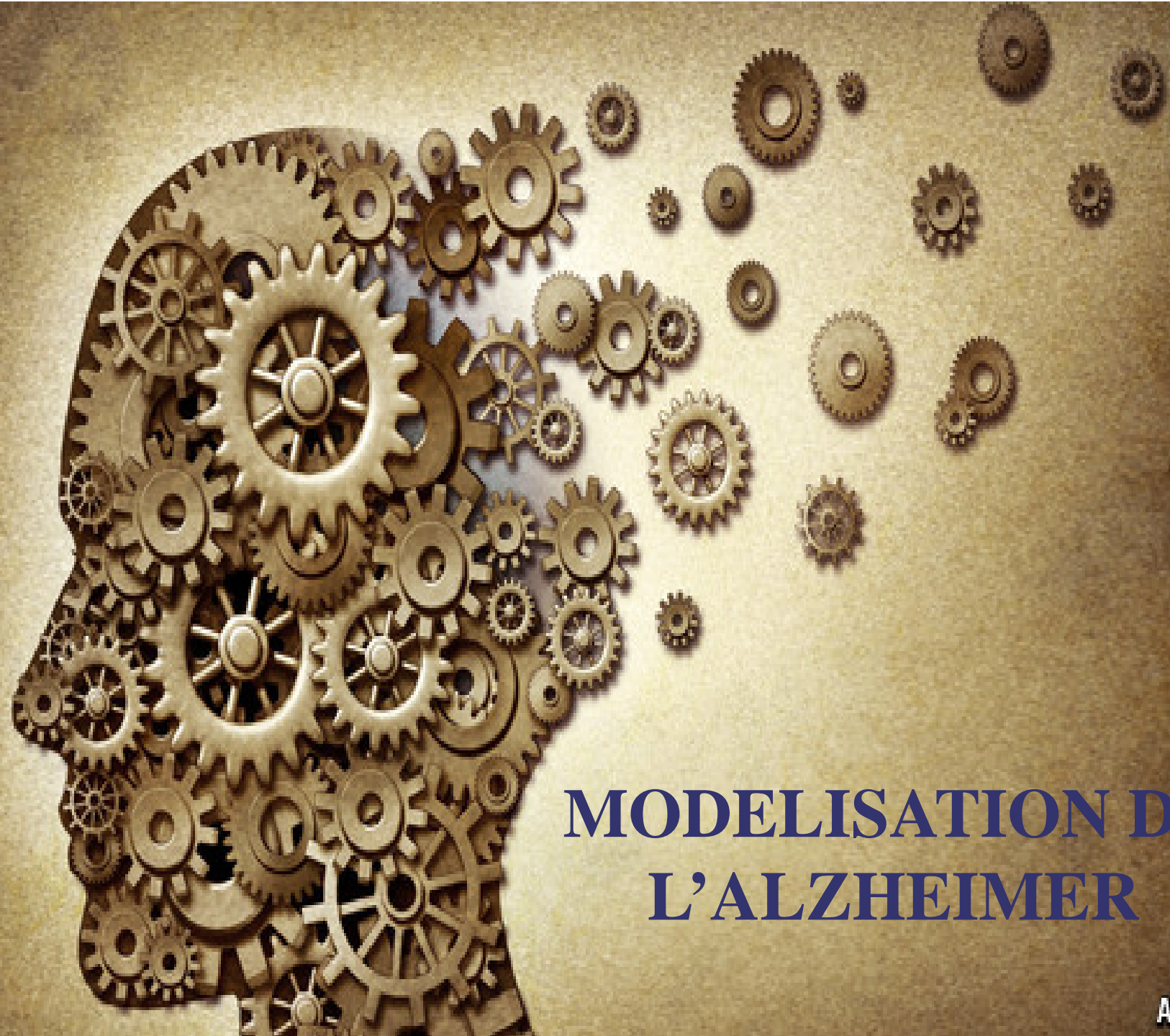
- Alzheimer et Diabète: maladies multidimensionnelles et progressives au cours du temps
- Alzheimer: Perte progressive de la mémoire et des fonctions cognitives, altération des paramètres IRM (exemple: diminution du volume de l'hippocampe)
- Diabète: Titration régulière de l'insuline basée sur le glucose à jeun, contrôle de l'HbA1c (ou hémoglobine glyquée) afin de juger l'équilibre de la glycémie au cours des 2 à 3 mois qui précèdent un dosage sanguin, et répétition des épisodes d'hypoglycémie
- Objectifs:
 - Modéliser l'évolution conjointe au cours du temps de plusieurs critères d'évaluation
 - Caractériser les facteurs de l'évolution globale des maladies (exemple: dégradation lente versus rapide)
 - Prendre en compte la dépendance entre critères
 - Aborder quelques questions cliniques



Quels sont les facteurs contribuant le plus à cette dégradation?

Existe-t-il une corrélation entre le HbA1c et le nombre d'occurrence de l'hypoglycémie?
La corrélation entre les critères réduit-elle l'estimation de l'effet traitement?



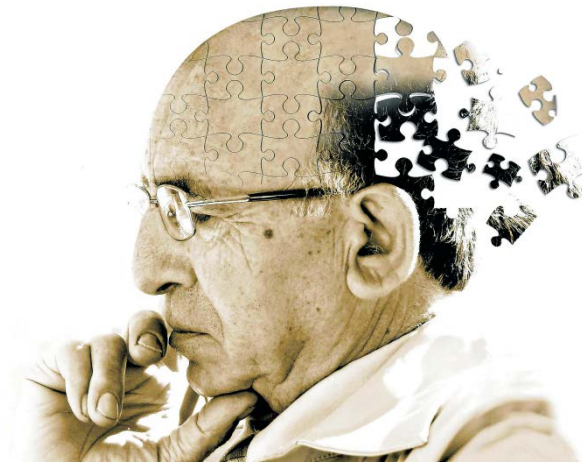


MODELISATION DE L'ALZHEIMER

MODELISATION DE L'ALZHEIMER

Modèle d'intérêt: le modèle multivarié à variable latente

- Construction du modèle avec $Y_{ijt} = \begin{cases} Y_{i1t} \\ Y_{i2t} \end{cases} = \begin{cases} CDR \\ Hippocampe \end{cases}$
- Définir les lois a priori
- Calcul des lois conditionnelles complètes
- Intégration des covariables (âge, genre, MMSE, etc...)
- Convergence, prédictions et évolution



MODELISATION DE L'ALZHEIMER

Modèle:

$$Y_{ijt} = \gamma_{ij}^{(1)} + \gamma_j^{(2)} \times L_{it} + \varepsilon_{ijt}^Y$$

$$[\varepsilon_{ijt}^Y] \sim \mathcal{N}(0, \sigma_j^2) \quad \forall i, j, t$$

$$L_{it} = \lambda_i \times t + \eta_i + \varepsilon_{it}^L \quad [L_{it}] \sim \mathcal{N}(X_{it}\beta_i, 1) \quad \text{with } X_{it} = (1, t) \text{ and } \beta_i = (\eta_i, \lambda_i)^T$$

$$[\varepsilon_{it}^L] \sim (0, 1) \quad \forall i, t$$

$$[\lambda_i | \theta_\lambda, \sigma_\lambda^2] \sim \mathcal{N}(\theta_\lambda, \sigma_\lambda^2)$$

$$[\eta_i | \theta_\eta, \sigma_\eta^2] \sim \mathcal{N}(\theta_\eta, \sigma_\eta^2)$$

$$[\gamma_{ij}^{(1)} | \theta_j^{(1)}, \sigma_j^{(1)^2}] \sim \mathcal{N}(\theta_j^{(1)}, \sigma_j^{(1)^2})$$

- i represents the patient number
- t represents time
- j represents the variable number ($j=1$ for CDR and $j=2$ for hippocampus)
- J_{ij} represents the number of visits where the patient has at least a given data of CDR or Hippocampus.

MODELISATION DE L'ALZHEIMER

Lois a priori: (motivées par des aspects de calculabilité et par manque d'informations)

- ✓ Utilisation de lois a priori peu informatives
- ✓ Utilisation des lois a priori conjuguées pour faciliter l'échantillonnage de Gibbs
- $[\sigma_j^2] \sim \text{inv-Gamma}(a,b)$
- $[\gamma_j^{(2)}] \sim \mathcal{N}(\theta_j^{\gamma(2)}, \sigma_j^{\gamma(2)^2}) \quad \forall j$
- $[\theta_\lambda] \sim \mathcal{N}(\mu_{\theta_\lambda}, \sigma_{\theta_\lambda}^2)$
- $[\theta_\eta] \sim \mathcal{N}(\mu_{\theta_\eta}, \sigma_{\theta_\eta}^2)$
- $[\sigma_\lambda^2] \sim \text{inv-Gamma}(a,b)$
- $[\sigma_\eta^2] \sim \text{inv-Gamma}(a,b)$
- $[\theta_j^{\gamma(1)}] \sim \mathcal{N}(\mu_{\theta_j^{\gamma(1)}}, \sigma_{\theta_j^{\gamma(1)}}^2) \quad \forall j$
- $[\sigma_j^{\gamma(1)}] \sim \text{inv-Gamma}(a,b) \quad \forall j$

MODELISATION DE L'ALZHEIMER

Lois conditionnelles complètes:

Appliquer le théorème de Bayes: $P(\theta|Y) \propto P(Y|\theta)P(\theta)$ où Y représente les données observées et θ un ensemble de paramètres, ignorer les paramètres non dépendants du paramètre d'intérêt.

● Exemple de la loi conditionnelle complète de λ_i

$$\begin{aligned} [\lambda_i | L_{it}, \theta_\lambda, \eta_i, \sigma_\lambda^2] &\propto \prod_{t=1}^{J_{ij}} [L_{it} | \lambda_i, \eta_i] [\lambda_i | \theta_\lambda, \sigma_\lambda^2] \\ &\propto \prod_{t=1}^{J_{ij}} \exp\left(-\frac{1}{2}(L_{it} - \lambda_i \times t - \eta_i)^T (L_{it} - \lambda_i \times t - \eta_i)\right) \\ &\quad \times \exp\left(-\frac{1}{2}(\lambda_i - \theta_\lambda)^T \frac{1}{\sigma_\lambda^2} (\lambda_i - \theta_\lambda)\right) \\ &\propto \exp\left(-\frac{1}{2}\left\{ \lambda_i^T \left(\sum_{t=1}^{J_{ij}} t^2 + \frac{1}{\sigma_\lambda^2} \right) \lambda_i - \lambda_i \left(\sum_{t=1}^{J_{ij}} t \times (L_{it} - \eta_i) + \frac{\theta_\lambda}{\sigma_\lambda^2} \right) \right. \right. \\ &\quad \left. \left. - \left(\sum_{t=1}^{J_{ij}} t \times (L_{it} - \eta_i) + \frac{\theta_\lambda}{\sigma_\lambda^2} \right) \lambda_i + \sum_{t=1}^{J_{ij}} (L_{it} - \eta_i)^2 + \frac{\theta_\lambda^2}{\sigma_\lambda^2} \right\} \right) \end{aligned}$$

$$[\lambda_i | L_{it}, \theta_\lambda, \eta_i, \sigma_\lambda^2] \sim \mathcal{N}(m_{\lambda_i}, s_{\lambda_i}^2)$$

$$\text{avec } s_{\lambda_i}^2 = \left(\sum_{t=1}^{J_{ij}} t^2 + \frac{1}{\sigma_\lambda^2} \right)^{-1} \text{ et } m_{\lambda_i} = s_{\lambda_i}^2 \times \left(\sum_{t=1}^{J_{ij}} t \times (L_{it} - \eta_i) + \frac{\theta_\lambda}{\sigma_\lambda^2} \right)$$

MODELISATION DE L'ALZHEIMER

Echantillonneur de Gibbs: (dans notre cas)

Phase 1: Initialisation

$$L_{it}^{[1]}, \gamma_{ij}^{(1)[1]}, \gamma_j^{(2)[1]}, \sigma_j^{2[1]}, \lambda_i^{[1]}, \eta_i^{[1]}, \theta_\lambda^{[1]}, \theta_\eta^{[1]}, \sigma_\lambda^{2[1]}, \sigma_\eta^{2[1]}, \theta_j^{\gamma(1)[1]}, \sigma_j^{\gamma(1)2[1]}$$

Phase 2: Itérations

$$L_{it}^{[niter+1]} \sim [L_{it} | Y_{ijt}, \lambda_i^{[niter]}, \eta_i^{[niter]}, \sigma_j^{2[niter]}]$$

$$\gamma_{ij}^{(1)[niter+1]} \sim [\gamma_{ij}^{(1)} | Y_{ijt}, \gamma_j^{(2)[niter]}, L_{it}^{[niter+1]}, \sigma_j^{2[niter]}]$$

$$\gamma_j^{(2)[niter+1]} \sim [\gamma_j^{(2)} | Y_{ijt}, \gamma_{ij}^{(1)[niter+1]}, L_{it}^{[niter+1]}, \sigma_j^{2[niter]}]$$

$$\sigma_j^{2[niter+1]} \sim [\sigma_j^2 | Y_{ijt}, \gamma_{ij}^{(1)[niter+1]}, \gamma_j^{(2)[niter+1]}, L_{it}^{[niter+1]}]$$

$$\lambda_i^{[niter+1]} \sim [\lambda_i | L_{it}^{[niter+1]}, \eta_i^{[niter]}, \theta_\lambda^{[niter]}, \sigma_\lambda^{2[niter]}]$$

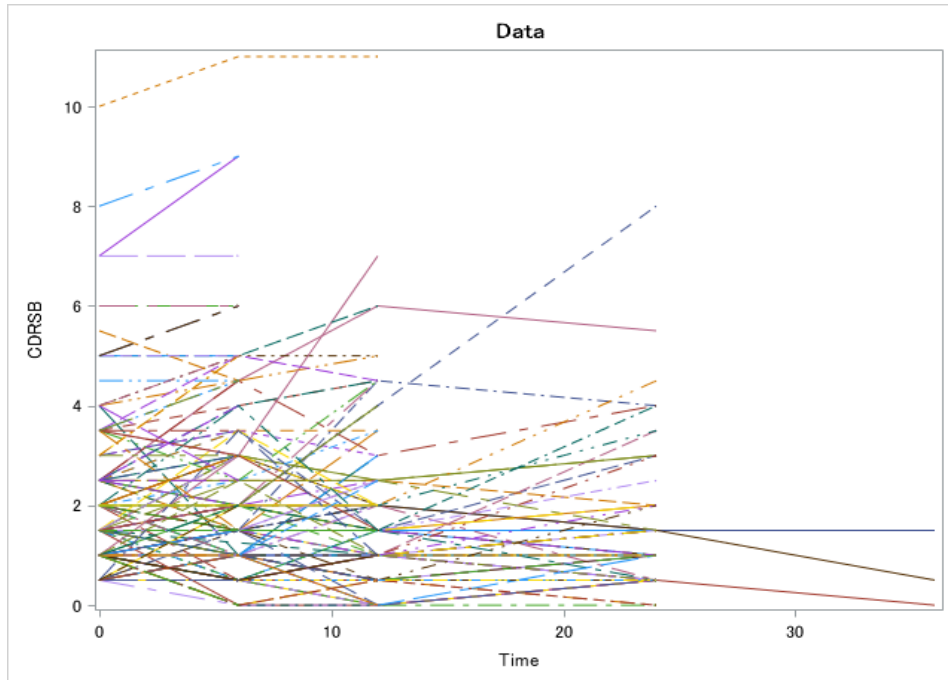
Phase 3: $i \leftarrow i + 1$ et aller en 2

MODELISATION DE L'ALZHEIMER

Description des données: (étude rétrospective, données manquantes)

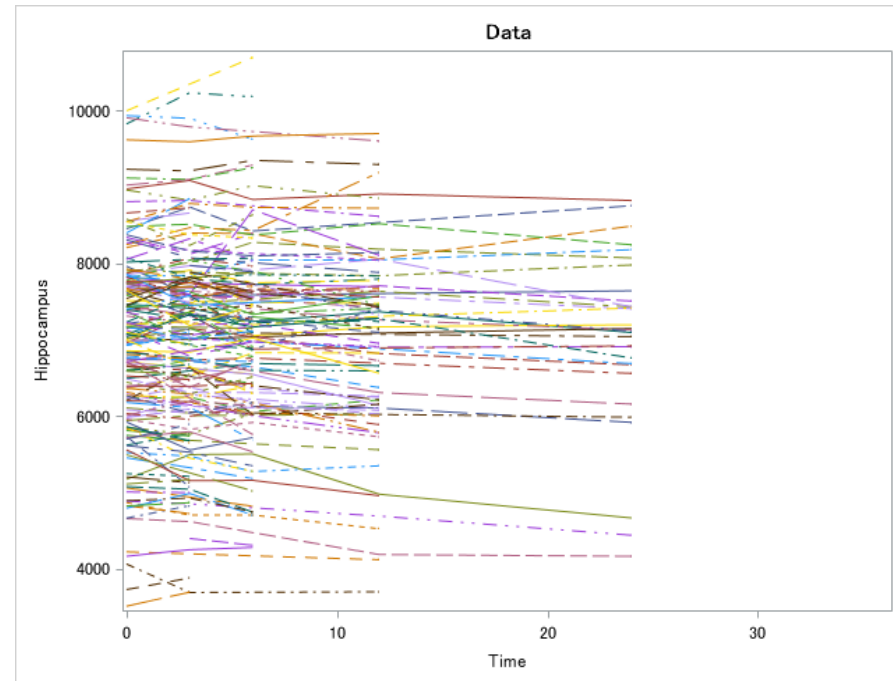
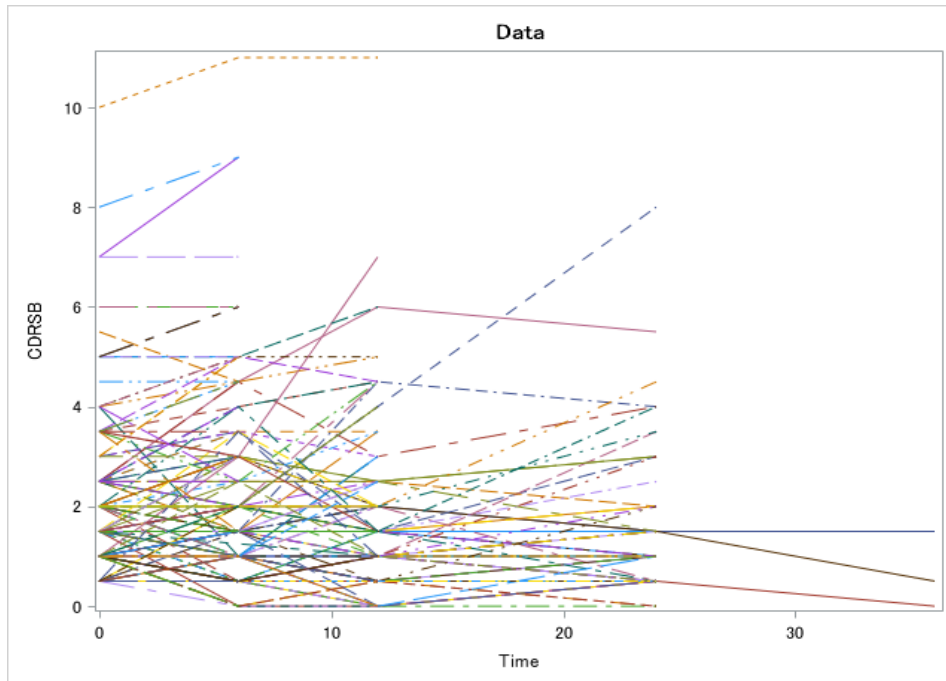
MODELISATION DE L'ALZHEIMER

Description des données: (étude rétrospective, données manquantes)



MODELISATION DE L'ALZHEIMER

Description des données: (étude rétrospective, données manquantes)



MODELISATION DE L'ALZHEIMER

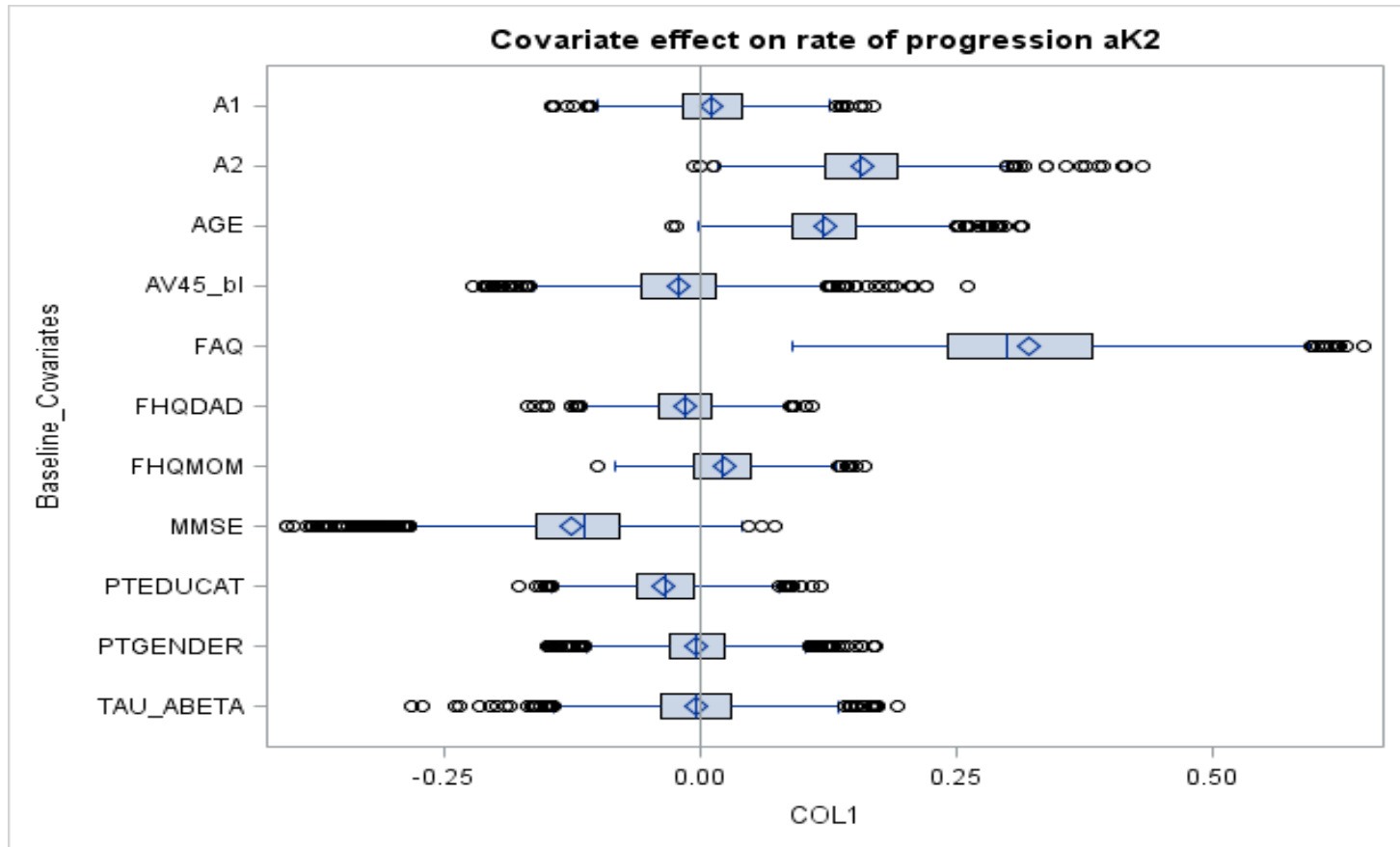
Convergence checks (Gelman and Rubin test) et visual checks:

Obs	Parameter	Between_chain	Within_chain	Estimate	UpperBound
1	theta_lambda	3.19	0.0057	1.4968	1.1117
2	sigma2_lambda	2.27	0.0083	1.2643	1.0694
3	theta_eta	1110.94	0.4251	2.7089	1.2318
4	sigma2_eta	401.65	1.8319	1.2772	1.1147
5	sigma2_1	0.01	0.0009	1.0117	1.0027

- w_{ij} le $i^{\text{ème}}$ ($i=1, \dots, n$) élément de la chaîne j ($j=1, \dots, m$)
- Variance inter-chaînes: $B = \frac{n}{m-1} \sum_{j=1}^m (\bar{\theta}_{.j} - \bar{\theta}_{..})^2$
- Variance intra-chaînes: $W = \frac{1}{m} \sum_{j=1}^m \left[\frac{1}{n-1} \sum_{i=1}^n (\bar{\theta}_{ij} - \bar{\theta}_{.j})^2 \right]$
- $\widehat{\text{Var}}(\boldsymbol{\theta}|\mathbf{y}) = \frac{n-1}{n} W + \frac{1}{n} B$
- Statistique de diagnostic: $\hat{R} = \sqrt{\frac{\widehat{\text{Var}}(\boldsymbol{\theta}|\mathbf{y})}{W}}$ qui doit être proche de 1

MODELISATION DE L'ALZHEIMER

Rapid Progressors (patients susceptibles de se dégrader rapidement):





MODELISATION DU DIABETE

MODELISATION DU DIABETE

Modèle: Traitements Lantus Vs Toujeo

- $Y_{i1t} \sim N\left(\gamma_{i1}^{(1)} + \gamma_1^{(2)} \times L_{it} + \beta_1 \times T_{it}, \sigma_1^2\right)$ où Y_{i1t} représente le HbA1c et β_1 représente l'effet du traitement sur Y_{i1t}
- $Y_{i2t} \sim \text{Pois}\left(\exp\left(\gamma_{i2}^{(1)} + \gamma_2^{(2)} \times L_{it} + \beta_2 \times T_{it}\right)\right)$ où Y_{i2t} représente le nombre d'occurrence de l'hypoglycémie et β_2 représente l'effet du traitement sur Y_{i2t}

T_{it} représente le traitement (0 pour Lantus et 1 pour Toujeo)

$$L_{it} \sim N(\eta_i + \lambda_i \times t, 1)$$

$$[\lambda_i \mid \theta_\lambda, \sigma_\lambda^2] \sim \mathcal{N}(\theta_\lambda, \sigma_\lambda^2)$$

$$[\eta_i \mid \theta_\eta, \sigma_\eta^2] \sim \mathcal{N}(\theta_\eta, \sigma_\eta^2)$$

$$[\gamma_{ij}^{(1)} \mid \theta_j^{\gamma(1)}, \sigma_j^{\gamma(1)^2}] \sim \mathcal{N}(\theta_j^{\gamma(1)}, \sigma_j^{\gamma(1)^2})$$

Lois a priori: Les lois a priori sont identiques et peu informatives comme dans le cas du modèle précédent.

MODELISATION DU DIABETE

Algorithme de Metropolis-Hastings:

- Lois conditionnelles complètes non usuelles
- Y_{i2t} suit une loi de poisson
- Utilisation de l'algorithme de Metropolis-Hastings

□ Phase 1: initialiser $x^{(0)}$

□ Phase 2: à l'itération i :

1. simuler $\tilde{x} \sim q(x|x^{i-1})$

2. calcul de $\alpha = \min\left\{1, \frac{\pi(\tilde{x})}{\pi(x^{i-1})} \frac{q(x^{i-1}|\tilde{x})}{q(\tilde{x}|x^{i-1})}\right\}$

3. accepter \tilde{x} avec la probabilité α :

$$x^{(i)} = \begin{cases} \tilde{x} & \text{avec la probabilité } \alpha \text{ (acceptation)} \\ x^{(i-1)} & \text{sinon (rejet)} \end{cases}$$

□ Phase 3: $i \leftarrow i + 1$ et aller en 2

MODELISATION DU DIABETE

Algorithme de Metropolis-Hastings (exemple de $\gamma_{i2}^{(1)}$):

❑ Phase 1: initialiser $\gamma_{i2}^{(1)(0)}$

❑ Phase 2: à l'itération i :

1. simuler $\widetilde{\gamma}_{i2}^{(1)} \sim q(\gamma_{i2}^{(1)} | \gamma_{i2}^{(1)i-1})$ (simuler la loi candidate $\widetilde{\gamma}_{i2}^{(1)}$ avec q une loi normale)

2. calcul de $\alpha = \min \left\{ 1, \frac{\pi(\widetilde{\gamma}_{i2}^{(1)})}{\pi(\gamma_{i2}^{(1)i-1})} \frac{q(\gamma_{i2}^{(1)i-1} | \widetilde{\gamma}_{i2}^{(1)})}{q(\widetilde{\gamma}_{i2}^{(1)} | \gamma_{i2}^{(1)i-1})} \right\}$ où $\pi(\widetilde{\gamma}_{i2}^{(1)})$ et $\pi(\gamma_{i2}^{(1)i-1})$ densités

respectives de $\widetilde{\gamma}_{i2}^{(1)}$ et $\gamma_{i2}^{(1)i-1}$ (le calcul des densités est détaillé dans les slides suivantes)

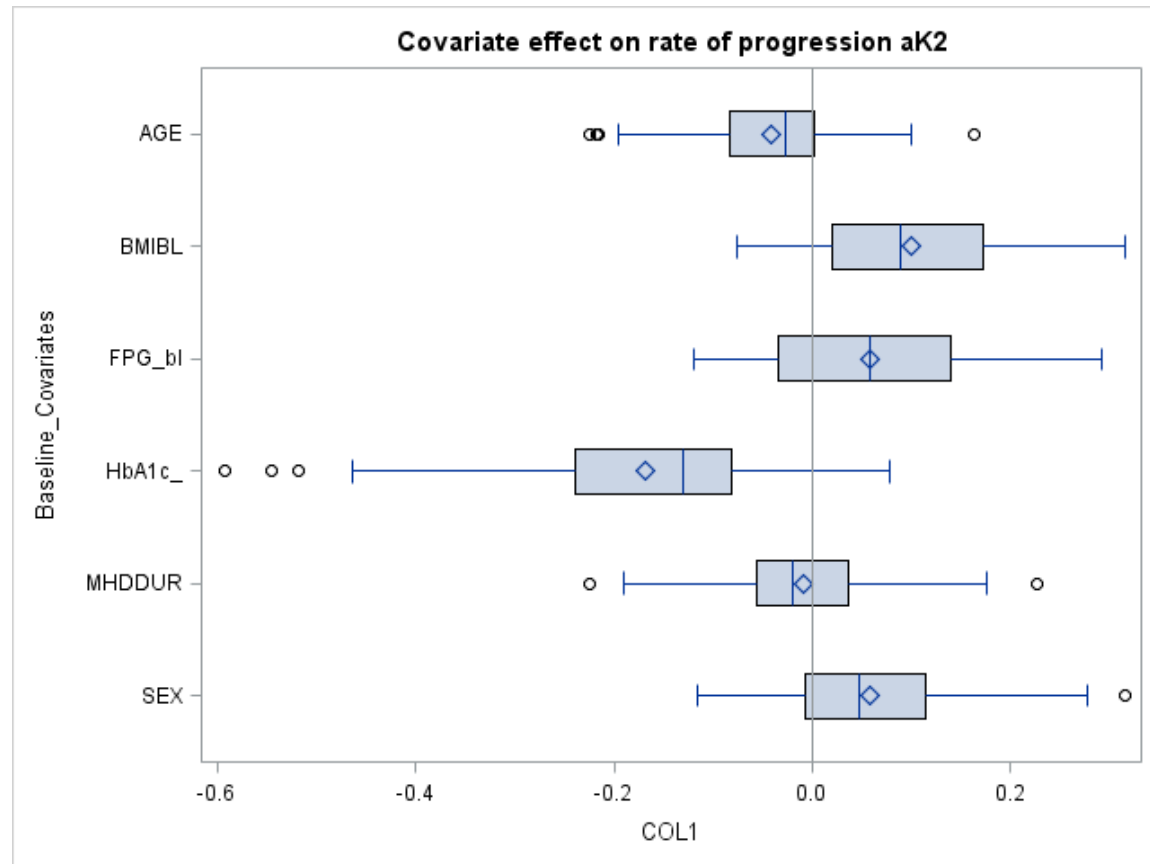
3. accepter $\widetilde{\gamma}_{i2}^{(1)}$ avec la probabilité α :

$$\gamma_{i2}^{(1)(i)} = \begin{cases} \widetilde{\gamma}_{i2}^{(1)} & \text{avec la probabilité } \alpha \text{ (acceptation)} \\ \gamma_{i2}^{(1)(i-1)} & \text{sinon (rejet)} \end{cases}$$

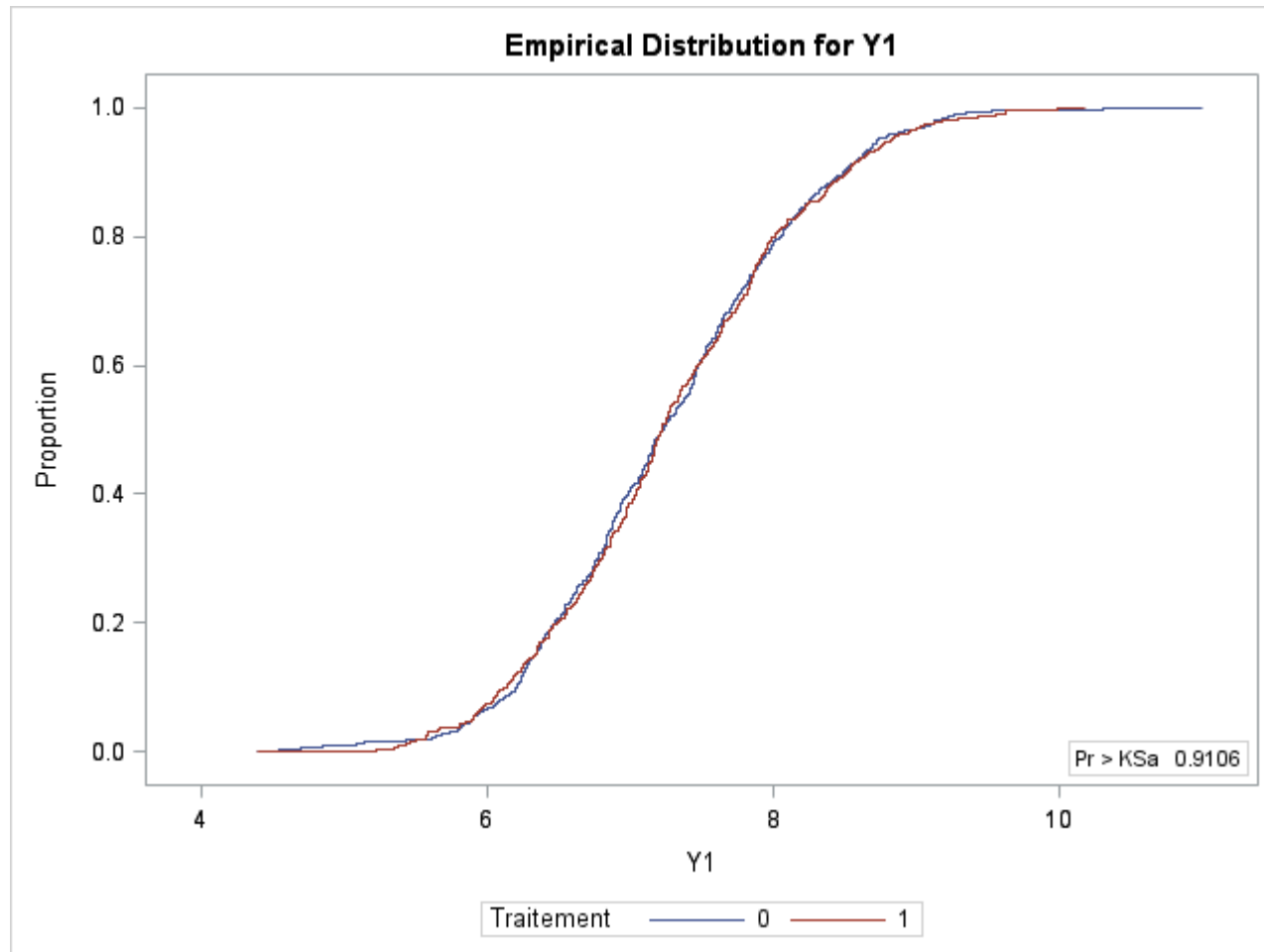
❑ Phase 3: $i \leftarrow i + 1$ et aller en 2

MODELISATION DU DIABETE

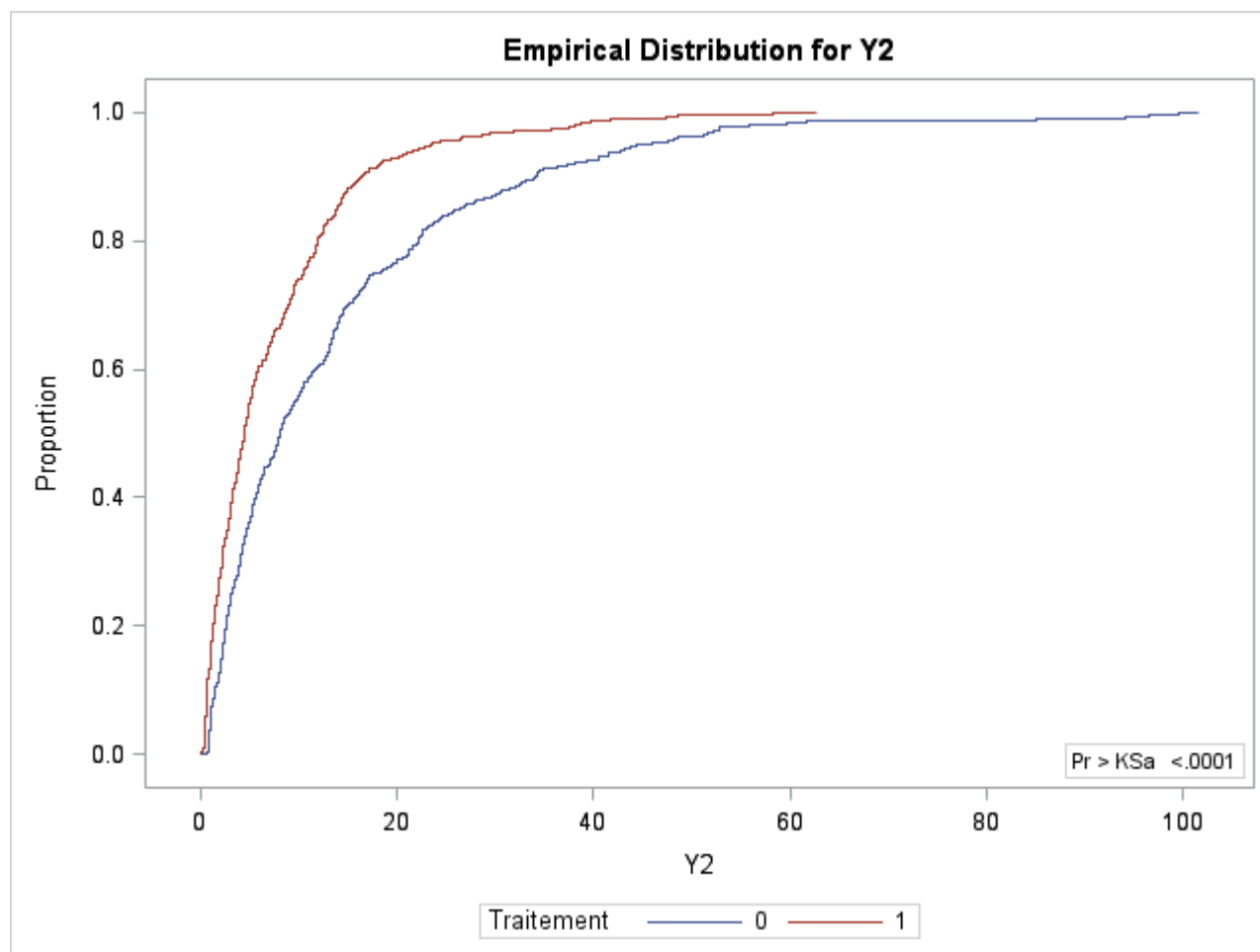
Effet des covariables sur l'évolution de la maladie (conjointement sur l'HbA1c et hypoglycémie):



MODELISATION DU DIABETE



MODELISATION DU DIABETE



CONCLUSION

CONCLUSION

- ✓ Modélisation via une variable latente prenant en compte des covariables

CONCLUSION

- ✓ Modélisation via une variable latente prenant en compte des covariables
- ✓ Application à Alzheimer et Diabète:
 - calcul des lois conditionnelles complètes et implémentation de l'algorithme de Gibbs et de Metropolis Hastings
 - vérification des convergences des paramètres du modèle
 - aborder des questions cliniques

CONCLUSION

- ✓ Modélisation via une variable latente prenant en compte des covariables
- ✓ Application à Alzheimer et Diabète:
 - calcul des lois conditionnelles complètes et implémentation de l'algorithme de Gibbs et de Metropolis Hastings
 - vérification des convergences des paramètres du modèle
 - aborder des questions cliniques
- ✓ Pistes et extensions envisageables pour améliorer le modèle:
 - mieux gérer les données manquantes
 - durée de l'étude et choix des patients
 - inclure d'autres variables ou covariables d'intérêt
 - intégrer une variable latente spécifique pour chaque variable

BIBLIOGRAPHIE

1. Casella, G. and George, E. (1992) “Explaining the Gibbs Sampler,” Amer. Statistician 46, 167-174
2. Flury, B. and Zoppe, A. (2000) “Exercises in EM,” Amer. Statistician 54, 207-209
3. Hastie, T., Tibshirani, R, and Friedman, J. The Elements of Statistical Learning. New York: Springer-Verlag, 2001, sections 8.5-8.6
4. Tierney, L. (1994) “Markov chains for exploring posterior distributions” (with discussion) Annals of Statistics 22, 1701-1762
5. Patrick Lam, convergence Diagnostics
6. Cécile Proust, Hélène Jacqmin-Gadda, Jeremy M. G. Taylor, Julien Ganiayre, and Daniel Commenges, A Nonlinear Model with Latent Process for Cognitive Evolution Using Multivariate Longitudinal Data
7. Sheng Luo, A Bayesian approach to joint analysis of multivariate longitudinal data and parametric accelerated failure time, Published online 6 September 2013 in Wiley Online library
8. Daniel G. Polhamus, Ph.D., James A. Rogers, Ph.D., William R. Gillespie, Ph.D., Jonathan French, Sc.D., Metrum Research Group, Tariffville, CT
9. Daniel G. Polhamus, James A. Rogers, William R. Gillespie, Jonathan French, Clinical Dementia Rating Modeling and Simulation : Joint progression of CDR and biomarkers in the ADNI cohort, Metrum research group, 2013
10. Ulrich Paquet, Bayesian inference for latent variable models, University of Cambridge, 2008
11. Nicolas Dobigeon, Modèles bayésiens hiérarchiques pour le traitement multi-capteur, Institut National Polytechnique de Toulouse, 2007

BIBLIOGRAPHIE

-
12. James A. Rogers, Longitudinal Joint Modeling of CDR items, ADAS-cog, and MRI Volumetrics, Metrum research group, 2014
 13. Pascal Minini, Bayesian Adaptive Design, Sanofi, 2011
 14. Michael W. Weiner MD <http://www.adni-info.org/>, 2005
 15. Pierrette Chagneau, Modélisation bayésienne hiérarchique pour la prédiction multivariée de processus ponctuels hétérogènes d'intensité liée à une variable prédite, Université de Montpellier, 2009
 16. Judith Rousseau, Statistique bayésienne, ENSAE Paris Tech, 2009
 17. Evans Gouno, Statistique Bayésienne, Université De Bretagne-sud, 2010
 18. Jean-Yves Tourneret, L'échantillonneur de Gibbs, Institut de recherche en informatique de Toulouse (IRIT) ENSEEIHT, 2006
 19. Jean-Yves Tourneret, Nicolas Dobigeon, Méthodes de Simulation, Institut de recherche en informatique de Toulouse (IRIT) ENSEEIHT, 2006
 20. Jean-Louis Foulley, Modèles bayésiens hiérarchiques/bases et applications en biostatistique, INRA, 2011
 21. B. Falissard, Comprendre et utiliser les statistiques dans les sciences de la vie, MASSON, 2005
 22. Marilyn S. Albert, Steven T. DeKosky, Dennis Dickson, Bruno Dubois, Howard H. Feldman, Nick C. Fox, Anthony Gamst, David M. Holtzman, William J. Jagust, Ronald C. Petersen, Peter J. Snyder, Maria C. Carrillo, Bill Thies, Creighton H. Phelps, The diagnosis of mild cognitive impairment due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimers Association workgroups on diagnostic guidelines for Alzheimer's disease, 2011
-

BIBLIOGRAPHIE

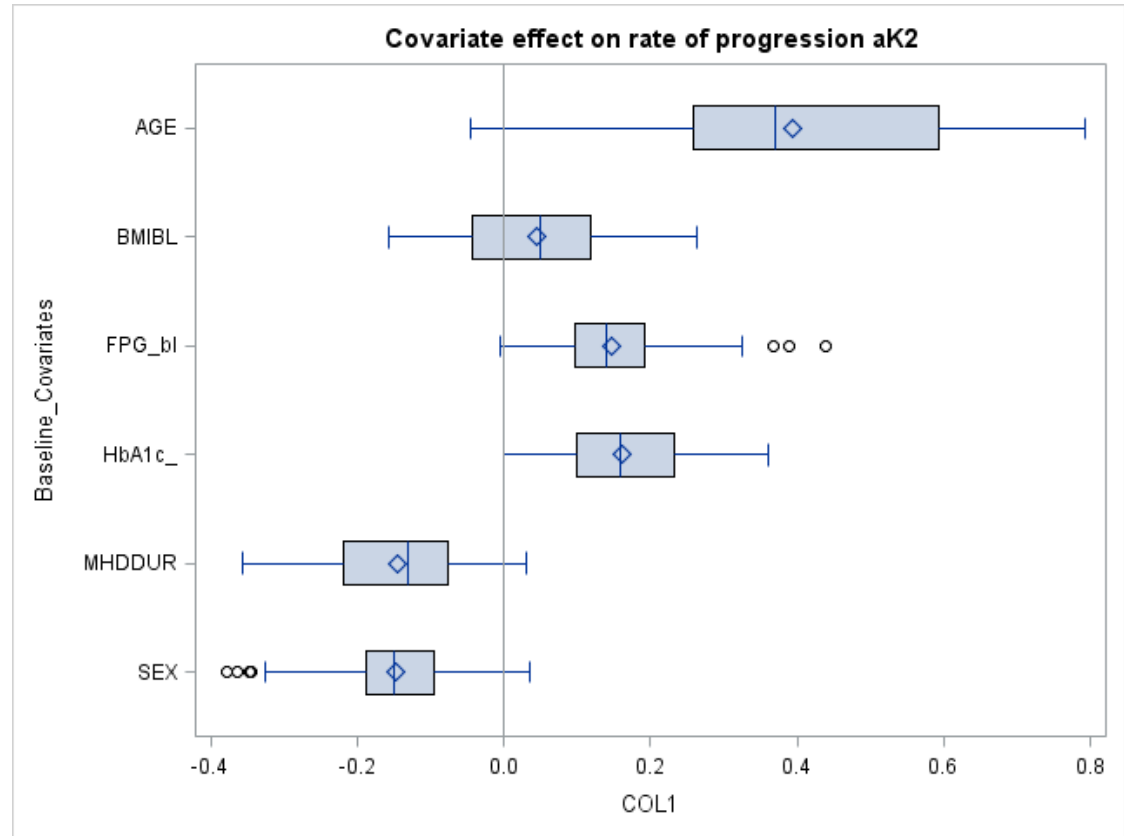
23. Geneviève Arsenault-Lapierre, Vieillissement et cognition, <http://geroscience.blogspot.fr/2011/02/cortisol-alzheimer101.html>, 2011
24. Jack et al, Hypothetical model of dynamic biomarkers of the Alzheimer's pathological cascade, *Lancet Neurol.*, 2010
25. Site du groupe : <http://www.sanofi.com/>
26. <http://www.francealzheimer.org/comprendre-maladie/chiffres/692>
27. Aurélie Millier, Pharmaco-économie, ISUP, 2013
28. Sheng Luo et Jue Wang, Bayesian hierarchical model for multiple repeated measures and survival data : an application to Parkinson's disease, *Statistics in Medicine*, 2014
29. Betsy Garofalo MD, Fundamentals of Clinical Trials Endpoints, Michigan Technology and Research Institute
30. Christophe Abraham, Notes de cours Master 2, Quelques modèles bayésiens, 23 Mai 2014
31. Andrew Gelman, John B. Carlin, Hal S. Stern and Donald B. Rubin, *Bayesian Data Analysis* (second edition)
32. Donald A. Berry, *A bayesian perspective*
33. W.R. Gilks, S. Richardson and D.J Spiegelhalter, *Markov Chain Monte Carlo in practice*

ANNEXES

DIABETES MODELING: FPG (at month 6) and Hypoglycemia (54 mg/dL severe or confirmed during 26-week on treatment period)

Covariates effect on the disease evolution (jointly on FPG and hypoglycemia):

These boxplots show the covariates effect on the disease evolution (jointly on FPG and hypoglycemia evolution). As expected, the fasting plasma glucose and BMI increase with the disease, the boxplot shows us that it has a positive significant impact on the evolution of the disease, as well as SEX, AGE, MHDDUR and HbA1c_bl which also have a significant impact on the evolution.



DIABETES MODELING: FPG (at month 6) and Hypoglycemia (54 mg/dL severe or confirmed during 26-week on treatment period)

Variable	N	Mean	Std Dev	Minimum	Maximum
beta1	100	1.0093185	0.7050622	0	4.3753548
beta2	100	-7.8793410	4.0799282	-13.8582462	0

$$H_0^{(1)} : \mu_{\beta_1} = 0$$

$$H_0^{(2)} : \mu_{\beta_2} = 0$$

By applying Wald test on each effect, we obtain:

$$W_1 = \frac{(1.0093185)^2}{(0.7050622)^2} = 2,048 < \chi_{(1)}^2 = 3,84 \rightarrow \text{we do not reject } H_0^{(1)}.$$

$$W_2 = \frac{(-7.8793410)^2}{(4.0799282)^2} = 3,729 < \chi_{(1)}^2 = 3,84 \rightarrow \text{we do not reject } H_0^{(2)}.$$

DIABETES MODELING: FPG (at month 6) and Hypoglycemia (54 mg/dL severe or confirmed during 26-week on treatment period)

Our essential goal was to check if there is a difference of treatments on each endpoint, by taking into consideration the correlation between FPG and occurrences of Hypoglycemia.

For β_1 , $H_0: \mu_{\beta_1} = 0$, we apply the PROC MEANS procedure in SAS. The Wald test (in the previous slide) and the p-value reveal that H_0 is not rejected (with a risk of 5%): No treatment difference on FPG.

For β_2 , $H_0: \mu_{\beta_2} = 0$. By applying the same test, we can acknowledge the non rejection of H_0 : No treatment difference on Hypoglycemia.

These results lead us to one conclusion: Considering the correlation between both endpoints, there is actually no difference of treatments on both endpoints.

DIABETES MODELING: FPG (at month 6) and Hypoglycemia (54 mg/dL severe or confirmed during 26-week on treatment period)

Y1 AND Y2 (PREDICTED)

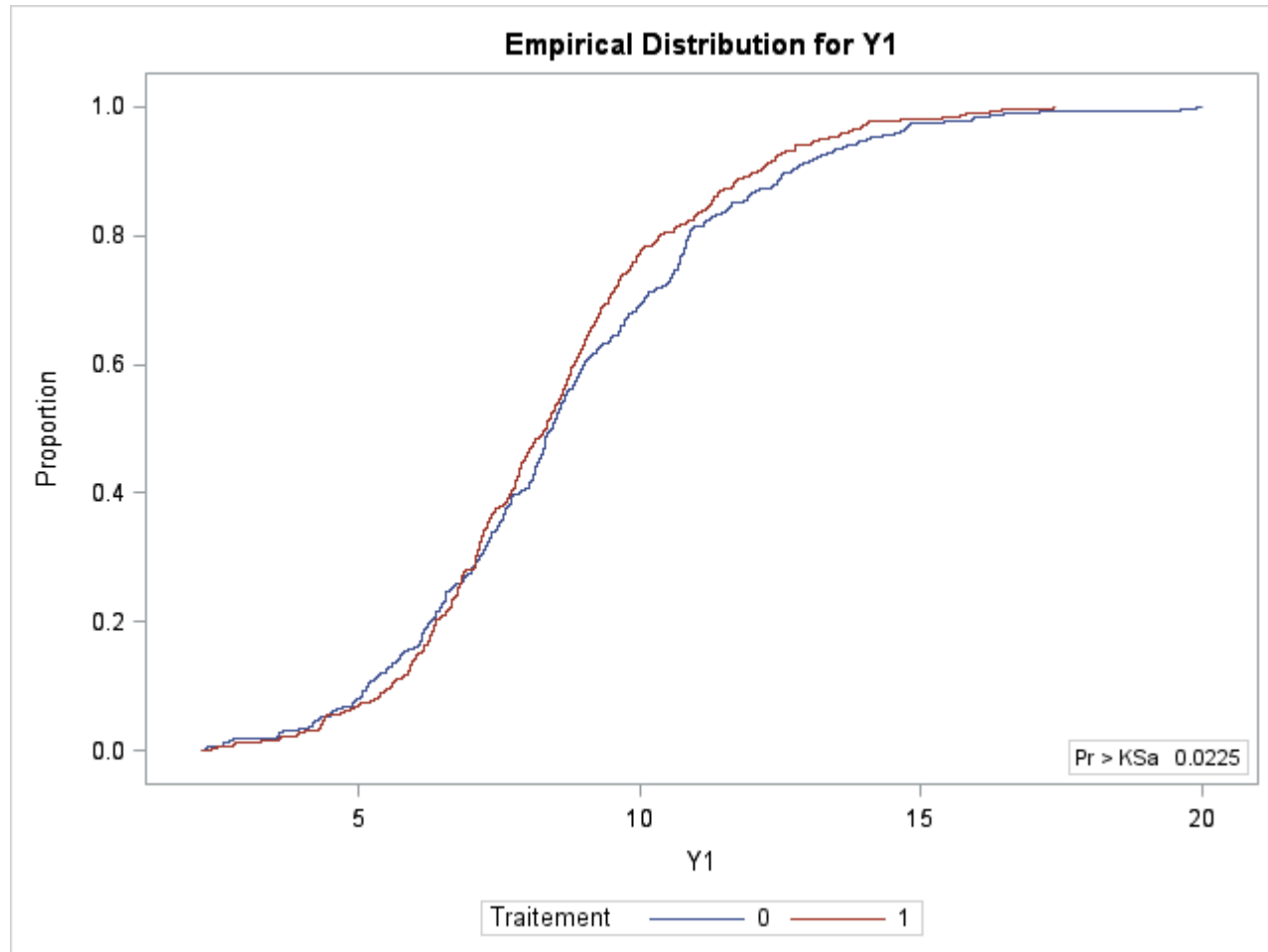
Traitement=0

Variable	N	Mean	Std Dev	Minimum	Maximum
Y1	670	8.7353381	2.9649028	2.1883972	20.0115444
Y2	670	13.7553954	16.0009374	0.5449622	114.1626941

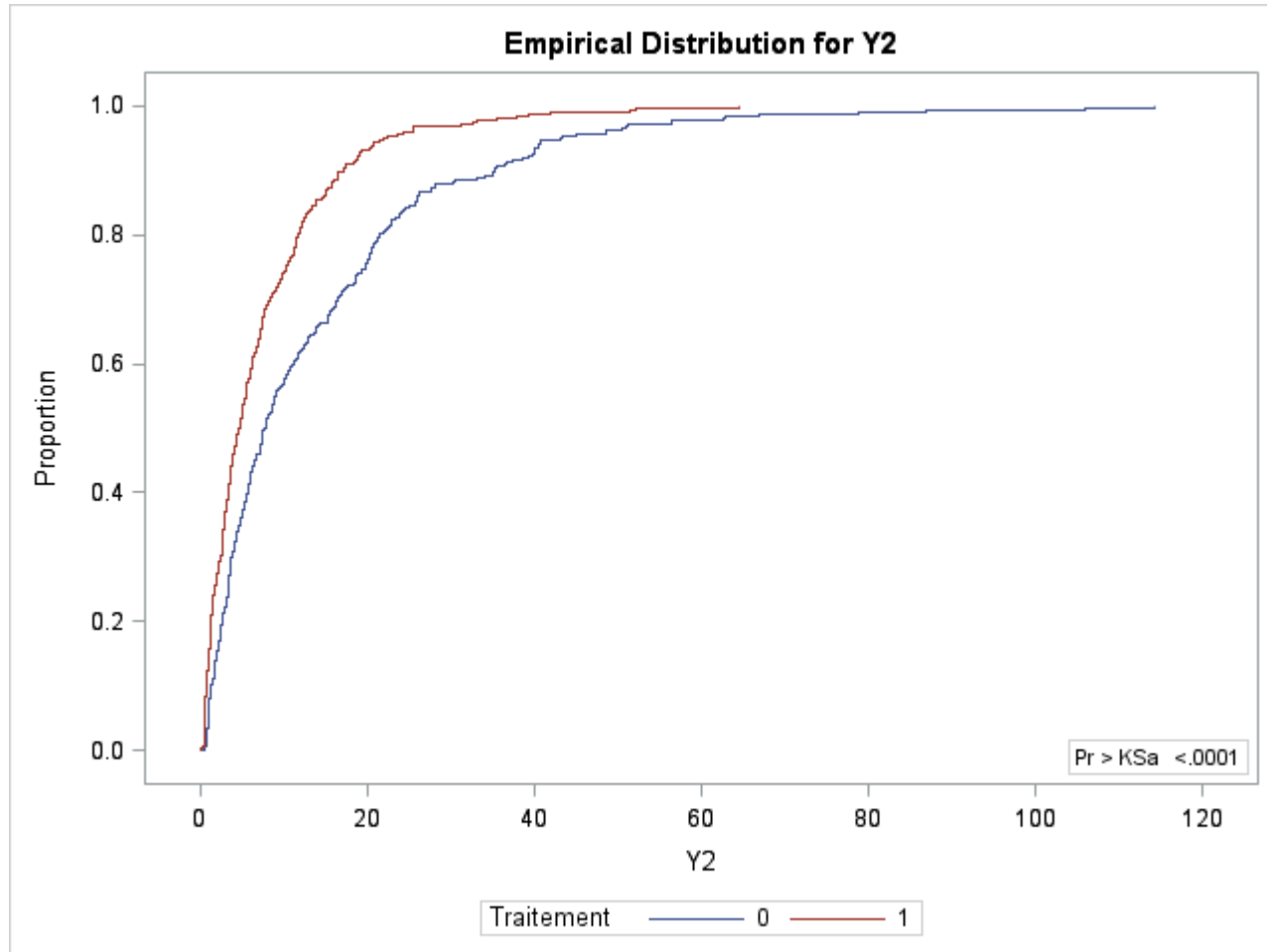
Traitement=1

Variable	N	Mean	Std Dev	Minimum	Maximum
Y1	614	8.4863234	2.5982848	2.2314622	17.3778549
Y2	614	7.4497041	8.7120741	0.0209246	64.4276623

DIABETES MODELING: FPG (at month 6) and Hypoglycemia (54 mg/dL severe or confirmed during 26-week on treatment period)



DIABETES MODELING: FPG (at month 6) and Hypoglycemia (54 mg/dL severe or confirmed during 26-week on treatment period)



DIABETES MODELING: FPG (at month 6) and Hypoglycemia (54 mg/dL severe or confirmed during 26-week on treatment period)

In order to be able to distinguish which treatment has more effectiveness and on which variable, we plotted these two graphs above (slides 16 and 17).

- The first graph represents the cumulative percentage of patients depending on FPG values and according to each treatment (the blue curve refers to Lantus and the red one refers to Toujeo). We can see that both curves are almost merged. The first graph demonstrates the fact that Toujeo and Lantus have at least the same potential to reduce FPG and therefore at least the same healing control for FPG.
- The second graph represents the cumulative percentage of patients depending on the number of hypoglycemia occurrences according to each treatment (blue curve for Lantus and red curve for Toujeo). We notice that the red curve is higher than the blue one, which confirms that there are more patients with less number of hypoglycemia occurrences using Toujeo treatment. So the second graph shows that Toujeo is more effective in reducing the number of hypoglycemia occurrences which is in favor of Toujeo.

MODELISATION DE L'ALZHEIMER

- **Aborder un modèle simple: le modèle bayésien probit**

Y_i = réponse binaire du patient i

X_i = variable prédictive du patient i

- Y est binaire et la variable latente est linéaire: $Z_i = X_i\beta + \varepsilon_i$, $X_i = \begin{pmatrix} 1 \\ x_i \end{pmatrix}^T$, $Y_i = I(Z_i \geq 0)$
- Calcul des lois conditionnelles complètes

- **Vers un modèle probit univarié catégoriel ordonné**

- Y est une variable catégorielle: $Y_i = \sum_{k=0}^{K-1} k \times I(c_k < Z_i \leq c_{k+1})$ where $c_0 = -\infty$ and $c_K = +\infty$
- Ajout d'une loi de seuils pour chaque catégorie k de la variable $c = (c_1, c_2, \dots, c_{K-1})$
- Calcul des nouvelles lois conditionnelles complètes
- Intégration des covariables (âge, genre, MMSE, etc...) et changement au niveau de la structure de la variable latente

MODELISATION DE L'ALZHEIMER

❖ Vers le modèle d'intérêt: le modèle multivarié à variable latente

➤ Construction du modèle avec $Y_{ijt} = \begin{cases} Y_{i1t} \\ Y_{i2t} \end{cases} = \begin{cases} CDR \\ Hippocampe \end{cases}$

➤ Lois a priori et lois conditionnelles complètes

➤ Convergence, prédictions et évolution

MODELISATION DE L'ALZHEIMER

Echantillonneur de Gibbs: (Algorithme général)

Phase 1: Initialisation

Initialiser $x^{[1]} = (x_1^{[1]}, \dots, x_N^{[1]})$

Phase 2: Itérations

A l'itération i : simuler

$$x_1^{[i+1]} \sim [\cdot | x_2^{(i)}, \dots, x_N^{(i)}]$$

$$x_2^{[i+1]} \sim [\cdot | x_1^{(i+1)}, x_3^{(i)}, \dots, x_N^{(i)}]$$

...

$$x_N^{[i+1]} \sim [\cdot | x_2^{(i+1)}, \dots, x_{N-1}^{(i+1)}]$$

Phase 3: $i \leftarrow i + 1$ et aller en 2

MODELISATION DE L'ALZHEIMER

Echantillonneur de Gibbs: (dans notre cas)

Phase 1: Initialisation

$$L_{it}^{[1]}, \gamma_{ij}^{(1)[1]}, \gamma_j^{(2)[1]}, \sigma_j^{2[1]}, \lambda_i^{[1]}, \eta_i^{[1]}, \theta_\lambda^{[1]}, \theta_\eta^{[1]}, \sigma_\lambda^{2[1]}, \sigma_\eta^{2[1]}, \theta_j^{\gamma(1)[1]}, \sigma_j^{\gamma(1)2[1]}$$

Phase 2: Itérations

$$L_{it}^{[niter+1]} \sim [L_{it} | Y_{ijt}, \lambda_i^{[niter]}, \eta_i^{[niter]}, \sigma_j^{2[niter]}]$$

$$\gamma_{ij}^{(1)[niter+1]} \sim [\gamma_{ij}^{(1)} | Y_{ijt}, \gamma_j^{(2)[niter]}, L_{it}^{[niter+1]}, \sigma_j^{2[niter]}]$$

$$\gamma_j^{(2)[niter+1]} \sim [\gamma_j^{(2)} | Y_{ijt}, \gamma_{ij}^{(1)[niter+1]}, L_{it}^{[niter+1]}, \sigma_j^{2[niter]}]$$

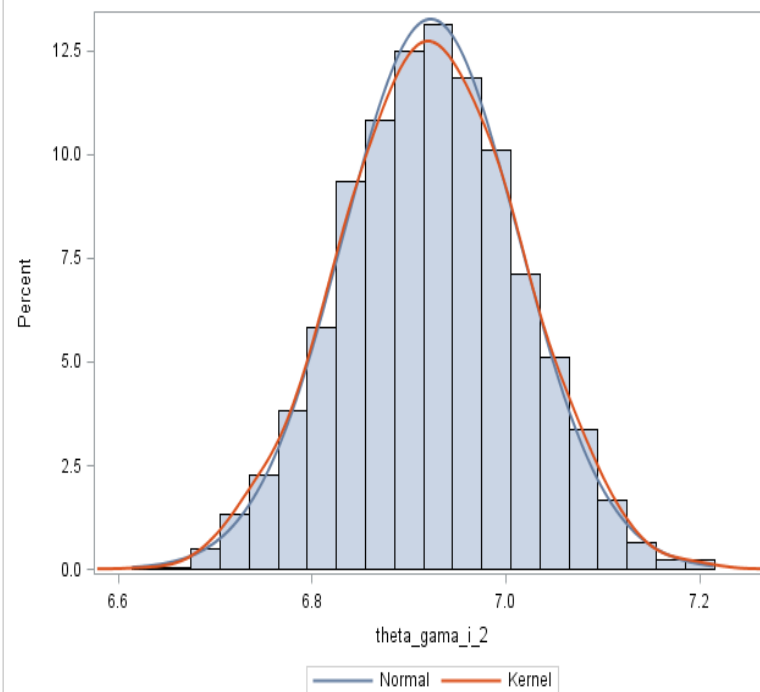
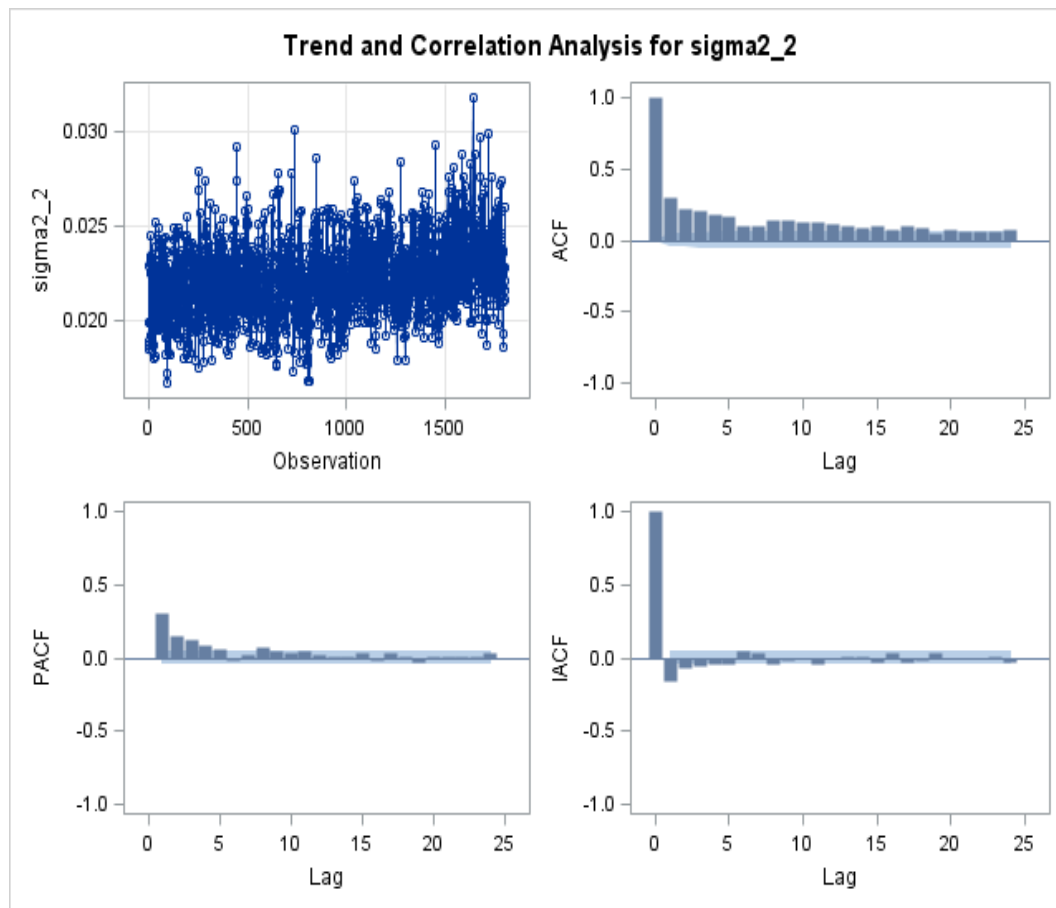
$$\sigma_j^{2[niter+1]} \sim [\sigma_j^2 | Y_{ijt}, \gamma_{ij}^{(1)[niter+1]}, \gamma_j^{(2)[niter+1]}, L_{it}^{[niter+1]}]$$

$$\lambda_i^{[niter+1]} \sim [\lambda_i | L_{it}^{[niter+1]}, \eta_i^{[niter]}, \theta_\lambda^{[niter]}, \sigma_\lambda^{2[niter]}]$$

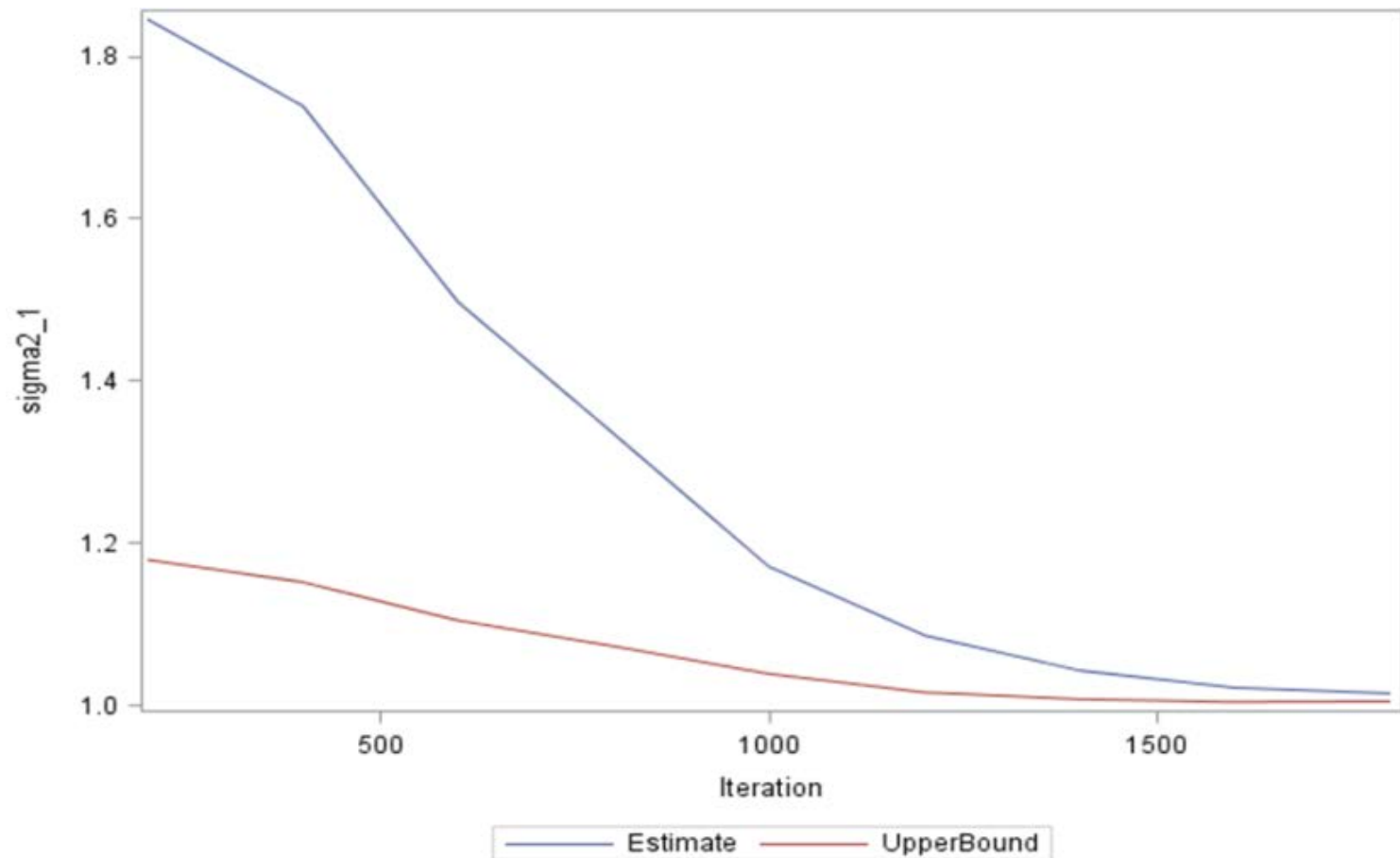
Phase 3: $i \leftarrow i + 1$ et aller en 2

MODELISATION DE L'ALZHEIMER

Traces of parameters:

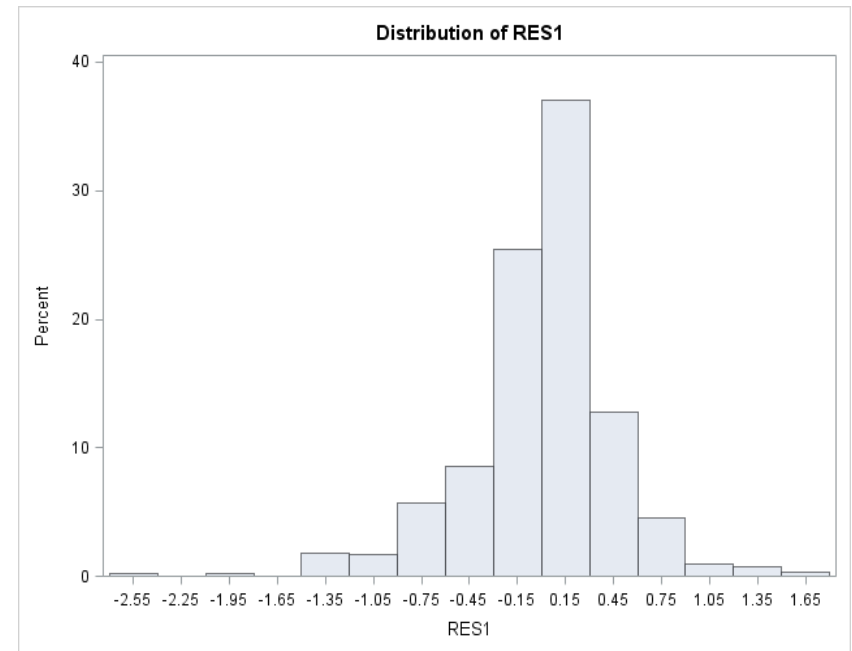
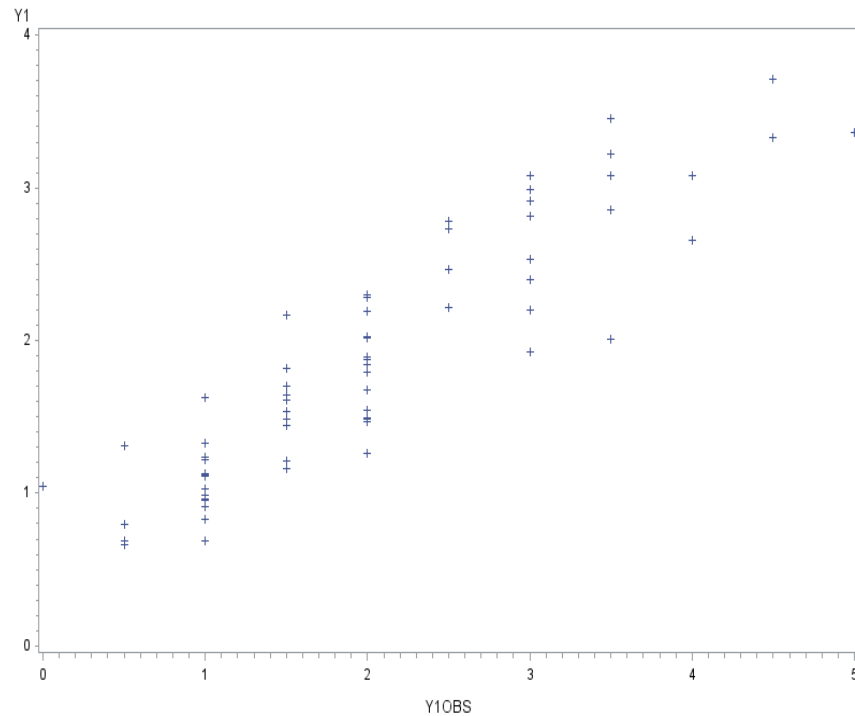


MODELISATION DE L'ALZHEIMER



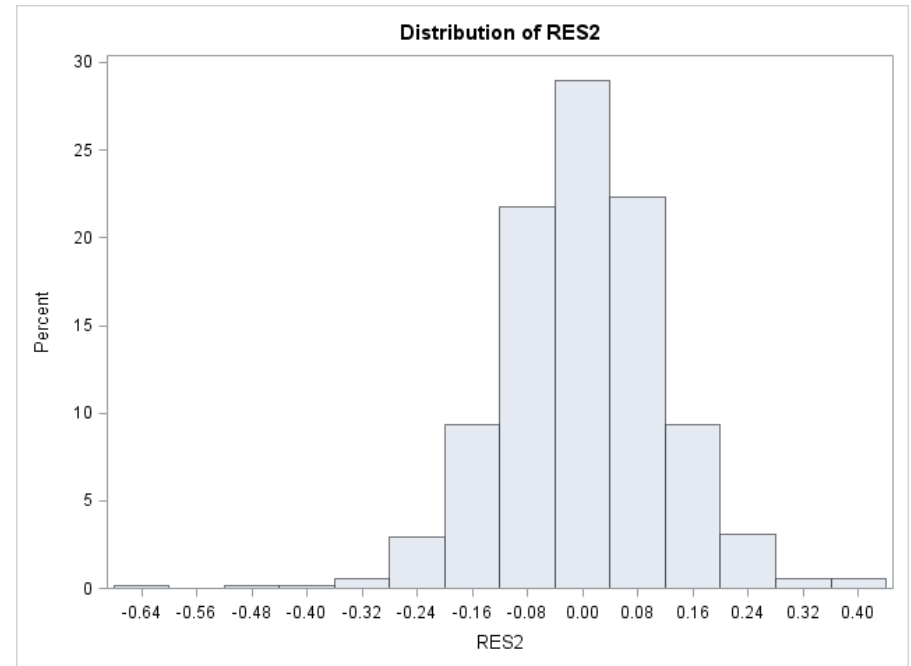
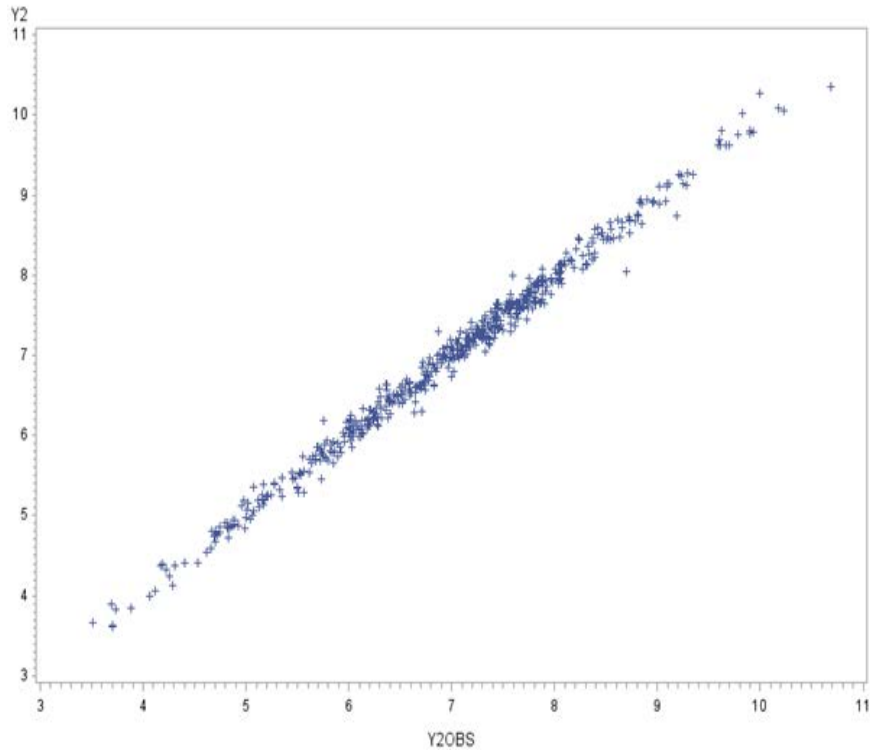
MODELISATION DE L'ALZHEIMER

Prévisions et résidus de Y1:



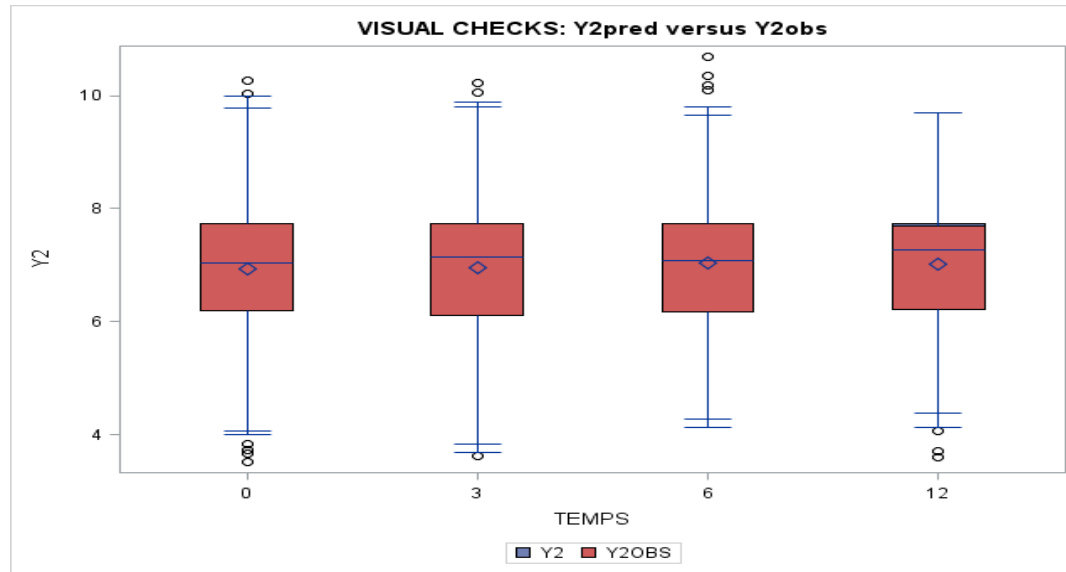
MODELISATION DE L'ALZHEIMER

Prévisions et résidus de Y2:



MODELISATION DE L'ALZHEIMER

Visual predictive checks pour Y2:



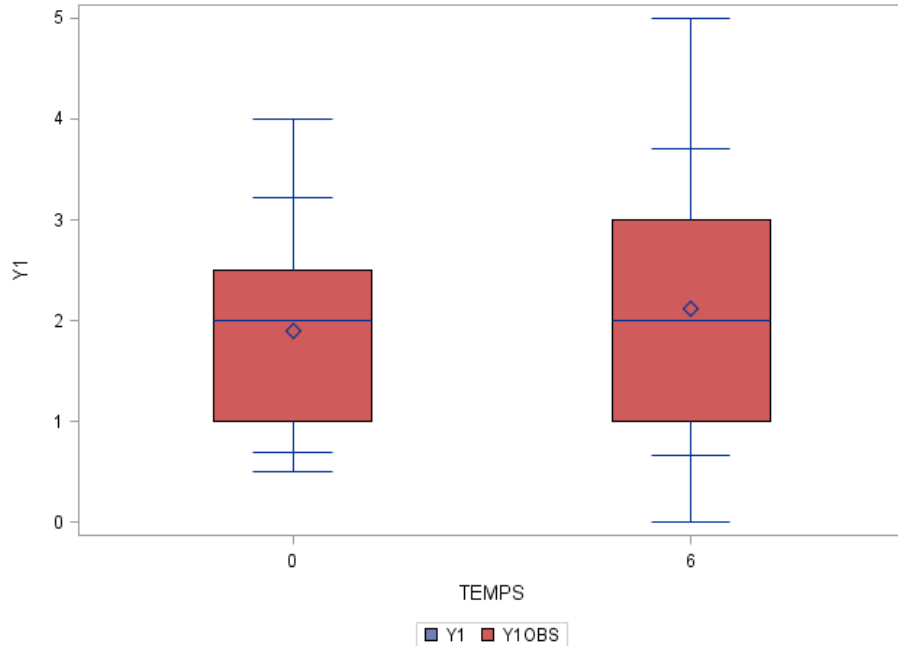
➡ Par profil: pour mieux visualiser l'évolution:

- profil 1 défini par les patients qui ont été suivis jusqu'à 6 mois
- profile 2 défini par les patients qui ont été suivis jusqu'à 12 mois
- profile 3 définie par les patients qui ont été suivis jusqu'à 18 mois
- profile 4 définie par les patients qui ont été suivis jusqu'à 24 mois
- profile 5 défini par les patients qui ont été suivis jusqu'à 30 mois
- profile 6 définie par les patients qui ont été suivis jusqu'à 36 moi

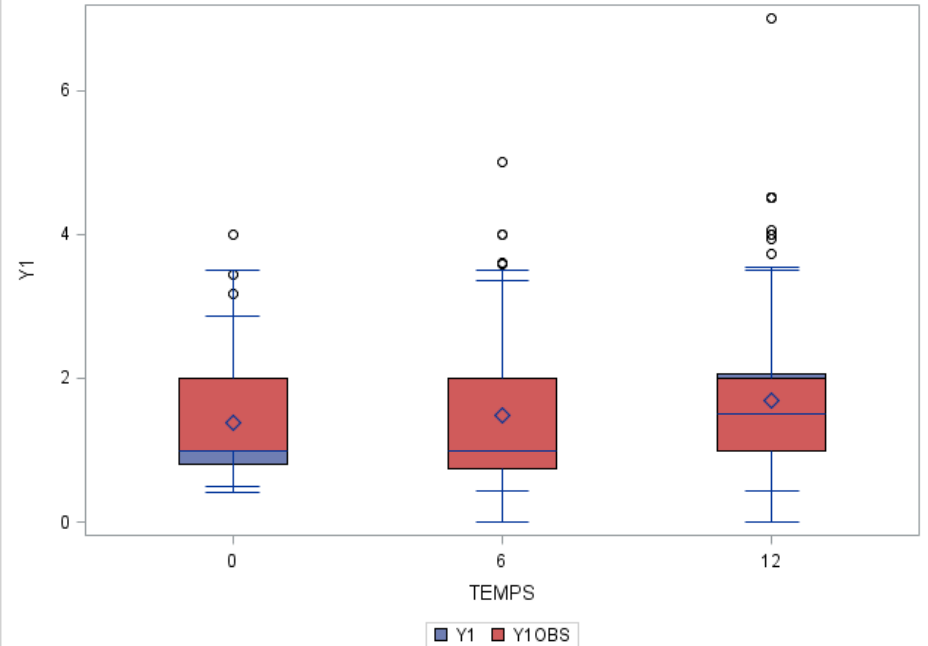
MODELISATION DE L'ALZHEIMER

Visual predictive checks pour Y1 par profil:

VISUAL CHECKS: Y1pred versus Y1obs BY PROFIL=1

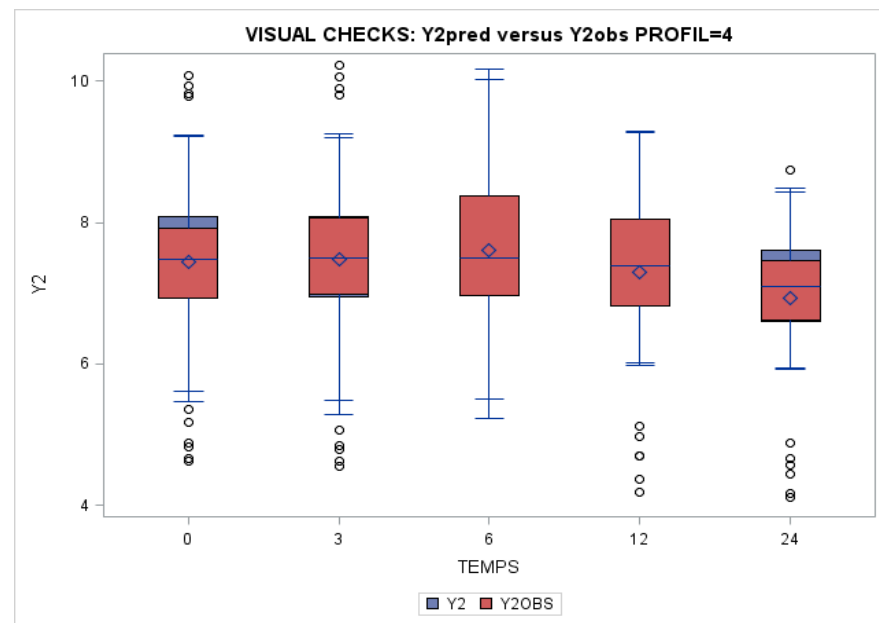
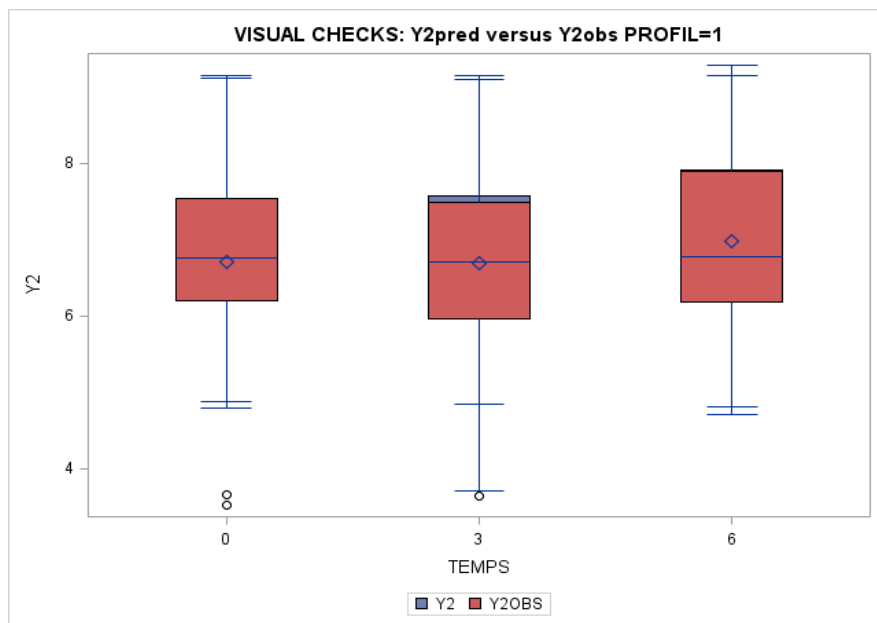


VISUAL CHECKS: Y1pred versus Y1obs PROFIL=2



MODELISATION DE L'ALZHEIMER

Visual predictive checks pour Y2 par profil:

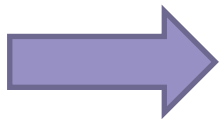


MODELISATION DE L'ALZHEIMER

Intégration des covariables:

On s'intéresse aux covariables suivantes: AGE, APGEN1, APGEN2, MMSE, GENRE, AV45_bl, FAQ, EDUCATION, FHQMOM, FHQDAD, TAU_ABETA

Notre objectif est d'évaluer l'effet de ces covariables non pas sur l'élargissement de l'état pathologique, mais sur la croissance de la maladie compte tenu du temps.



$$L_{it} = \eta_i + W_{ij}\alpha_1 + (\lambda_i + W_{ij}\alpha_2)X_{ij} + \varepsilon_{ij}$$

Seulement trois lois conditionnelles changent (L_{it} , η_i et λ_i) et deux nouvelles lois sont à calculer (α_1 : l'ajustement des covariables sur la constante et α_2 : l'ajustement des covariables sur la pente).

A noter que les covariables doivent être centrées et réduites afin de les rendre comparables.

MODELISATION DU DIABETE

Lois conditionnelles complètes identiques:

Le calcul des lois conditionnelles suivantes est identique:

$$\sigma_1^2 \quad \gamma_{i1}^{(1)} \quad \gamma_1^{(2)} \quad \lambda_i \quad \theta_\lambda \quad \sigma_\lambda^2 \quad \eta_i \quad \theta_\eta \quad \sigma_\eta^2 \quad \theta_j^{\gamma(1)} \quad \sigma_j^{\gamma(1)^2}$$

Lois conditionnelles complètes à recalculer:

- $[\sigma_2^2 | Y_{i2t}, \gamma_{i2}^{(1)}, \gamma_2^{(2)}, L_{it}]$ (paramètre à supprimer)
- $[\gamma_{i2}^{(1)} | Y_{i2t}, \gamma_2^{(2)}, L_{it}] \propto$

$$\prod_{t=1}^{J_{ij}} \exp \left(Y_{i2t} \left(\gamma_{i2}^{(1)} + \gamma_2^{(2)} L_{it} + \beta_2 \times T_{it} \right) - \exp \left(\gamma_{i2}^{(1)} + \gamma_2^{(2)} L_{it} + \beta_2 \times T_{it} \right) \right)$$

$$\times \exp \left(-\frac{1}{2} \left(\theta_2^{\gamma(1)} - \gamma_{i2}^{(1)} \right)^T \frac{1}{\sigma_2^{\gamma(1)^2}} \left(\theta_2^{\gamma(1)} - \gamma_{i2}^{(1)} \right) \right)$$

MODELISATION DU DIABETE

- $$\begin{aligned}
 & [L_{it} | \lambda_i, \eta_i, Y_{ijt}, \sigma_1^2] \propto \exp \left(-\frac{1}{2} (L_{it} - \eta_i - \lambda_i \times t)^T (L_{it} - \eta_i - \lambda_i \times t) \right) \\
 & \times \exp \left(-\frac{1}{2} \left(Y_{i1t} - \gamma_{i1}^{(1)} - \gamma_1^{(2)} L_{it} - \beta_1 \times T_{it} \right)^T \frac{1}{\sigma_1^2} \left(Y_{i1t} - \gamma_{i1}^{(1)} - \gamma_1^{(2)} L_{it} - \beta_1 \right. \right. \\
 & \times \left. \left. \frac{1}{Y_{(i2t)}!} \exp \left(Y_{i2t} \left(\gamma_{i2}^{(1)} + \gamma_2^{(2)} L_{it} + \beta_2 \times T_{it} \right) - \exp \left(\gamma_{i2}^{(1)} + \gamma_2^{(2)} L_{it} + \beta_2 \times T_{it} \right) \right) \right) \right)
 \end{aligned}$$
- $$\begin{aligned}
 & [\gamma_2^{(2)} | Y_{i2t}, \gamma_{i2}^{(1)}, L_{it}] \\
 & \propto \prod_{i=1}^N \prod_{t=1}^{J_{ij}} \exp \left(Y_{i2t} \left(\gamma_{i2}^{(1)} + \gamma_2^{(2)} L_{it} + \beta_2 \times T_{it} \right) - \exp \left(\gamma_{i2}^{(1)} + \gamma_2^{(2)} L_{it} + \beta_2 \times T_{it} \right) \right) \\
 & \times \exp \left(-\frac{1}{2} \left(\theta_2^{\gamma^{(2)}} - \gamma_2^{(2)} \right)^T \frac{1}{\left(\sigma_2^{\gamma^{(2)}} \right)^2} \left(\theta_2^{\gamma^{(2)}} - \gamma_2^{(2)} \right) \right)
 \end{aligned}$$

MODELISATION DU DIABETE

Nouvelles lois conditionnelles complètes: (des effets traitement sur chaque endpoint)

- $[\beta_1 | Y_{i1t}, \gamma_{i1}, \gamma_1, L_{it}] \propto \prod_{i=1}^N \prod_{t=1}^{J_{ij}} [Y_{i1t} | \gamma_{i1}, \gamma_1, L_{it}, \beta_1] [\beta_1 | \mu_{\beta_1}, \sigma_{\beta_1}^2]$
 $[\beta_1 | Y_{i1t}, \gamma_{i1}, \gamma_1, L_{it}] \sim \mathcal{N}(m_{\beta_1}, s_{\beta_1}^2)$
$$s_{\beta_1}^2 = \left(\frac{1}{\sigma_{\beta_1}^2} + \sum_{i=1}^N \sum_{t=1}^{J_{ij}} \frac{T_{it}}{\sigma_1^2} \right)^{-1}$$
$$m_{\beta_1} = s_{\beta_1}^2 \left(\frac{\mu_{\beta_1}}{\sigma_{\beta_1}^2} + \sum_{i=1}^N \sum_{t=1}^{J_{ij}} \frac{T_{it}}{\sigma_1^2} (Y_{i1t} - \gamma_{i1} - \gamma_1 L_{it}) \right)$$
- $[\beta_2 | Y_{i2t}, \gamma_{i2}, \gamma_2, L_{it}] \propto \prod_{i=1}^N \prod_{t=1}^{J_{ij}} [Y_{i2t} | \gamma_{i2}, \gamma_2, L_{it}, \beta_2] [\beta_2 | \mu_{\beta_2}, \sigma_{\beta_2}^2]$
 $\propto \prod_{i=1}^N \prod_{t=1}^{J_{ij}} \exp \left(Y_{i2t} \left(\gamma_{i2}^{(1)} + \gamma_2^{(2)} L_{it} + \beta_2 \times T_{it} \right) - \exp \left(\gamma_{i2}^{(1)} + \gamma_2^{(2)} L_{it} + \beta_2 \times T_{it} \right) \right)$
 $\times \exp \left(- \frac{1}{2} (\beta_2 - \mu_{\beta_2})^T \frac{1}{\sigma_{\beta_2}^2} (\beta_2 - \mu_{\beta_2}) \right)$

MODELISATION DU DIABETE

Effet des covariables:

$$L_{it} = \eta_i + W_{ij}\alpha_1 + (\lambda_i + W_{ij}\alpha_2)X_{ij} + \varepsilon_{ij}$$

Les lois conditionnelles complètes sont identiques, seul la loi de L_{it} change:

$$\begin{aligned} [L_{it} | \lambda_i, \eta_i, Y_{ijt}, \sigma_1^2] &\propto \exp\left(-\frac{1}{2}\left(L_{it} - (\eta_i + W_{ij}\alpha_1) - (\lambda_i + W_{ij}\alpha_2)X_{ij}\right)^T \right. \\ &\times \left(L_{it} - (\eta_i + W_{ij}\alpha_1) - (\lambda_i + W_{ij}\alpha_2)X_{ij}\right) \\ &\times \exp\left(-\frac{1}{2}\left(Y_{i1t} - \gamma_{i1}^{(1)} - \gamma_1^{(2)}L_{it} - \beta_1 \times T_{it}\right)^T \frac{1}{\sigma_1^2} \left(Y_{i1t} - \gamma_{i1}^{(1)} - \gamma_1^{(2)}L_{it} - \beta_1 \times T_{it}\right)\right) \\ &\times \frac{1}{Y_{(i2t)!}} \exp\left(Y_{i2t} \left(\gamma_{i2}^{(1)} + \gamma_2^{(2)}L_{it} + \beta_2 \times T_{it}\right) - \exp\left(\gamma_{i2}^{(1)} + \gamma_2^{(2)}L_{it} + \beta_2 \times T_{it}\right)\right) \end{aligned}$$



MODELISATION DU DIABETE

Relation entre HbA1c et le nombre d'occurrence d'hypoglycémie:

Pearson Correlation Coefficients Prob > r under H0: Rho=0 Number of Observations		
	Y1	Y2
Y1	1.00000 3659	- 0.59678 <.0001 1330
Y2	-0.59678 <.0001 1330	1.00000 1330

Pearson Correlation Coefficients Prob > r under H0: Rho=0 Number of Observations		
	Y1OBS	Y2OBS
Y1OBS	1.00000 3659	-0.57044 <.0001 1330
Y2OBS	-0.57044 <.0001 1330	1.00000 1330

MODELISATION DU DIABETE

Effets de traitement (Lantus vs Touejo):

Variable	N	Mean	Std Dev	Minimum	Maximum
beta1	100	0.3355743	0.6426280	-0.0585591	2.8845471
beta2	100	-7.0531795	4.8157434	-16.2338376	0

$H_0^{(1)} : \mu_{\beta_1} = 0$ (où β_1 est l'effet traitement sur HbA1c)

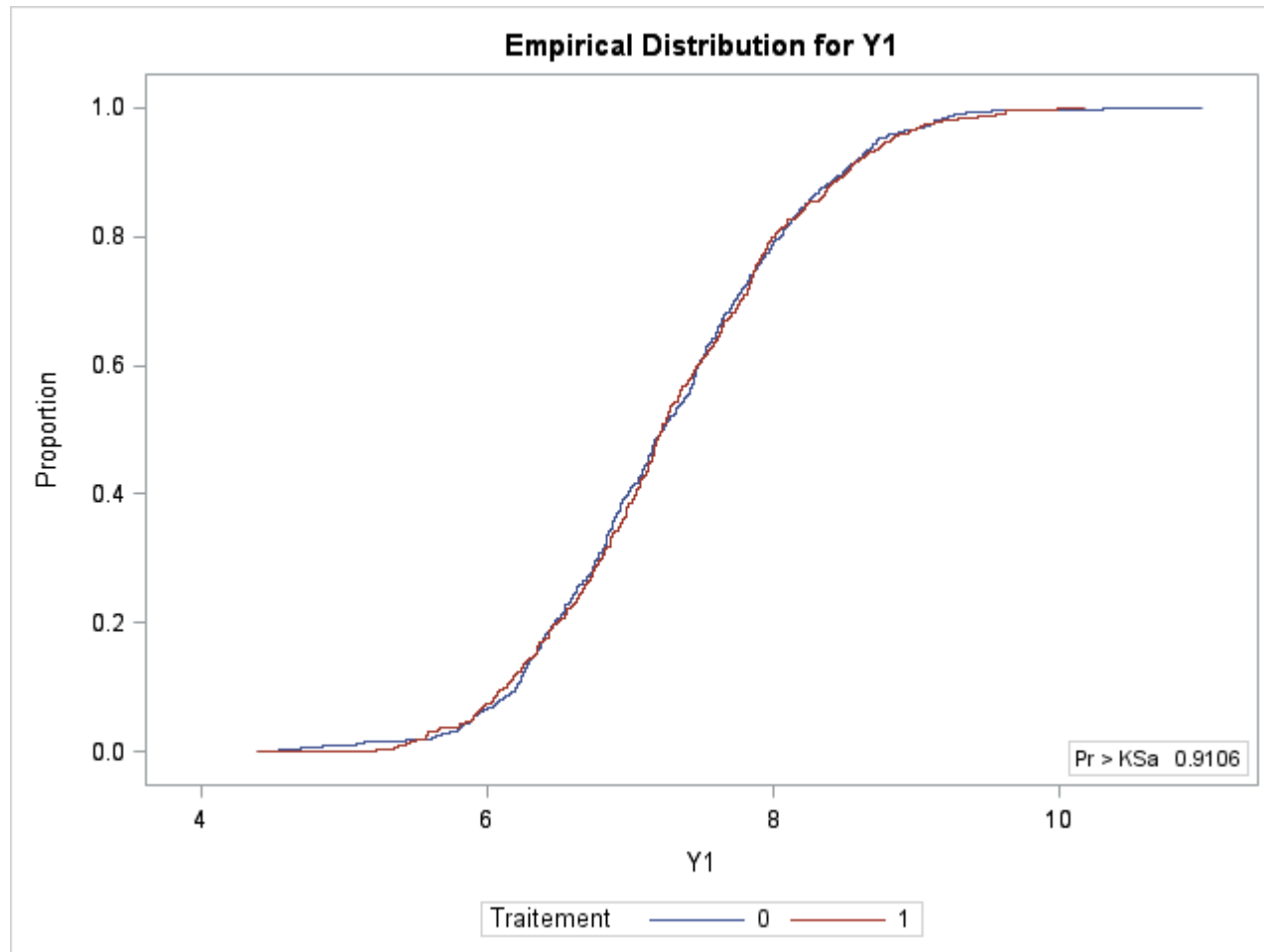
$H_0^{(2)} : \mu_{\beta_2} = 0$ (où β_2 est l'effet traitement sur les hypoglycémies)

En appliquant le test de Wald pour chaque effet, on obtient:

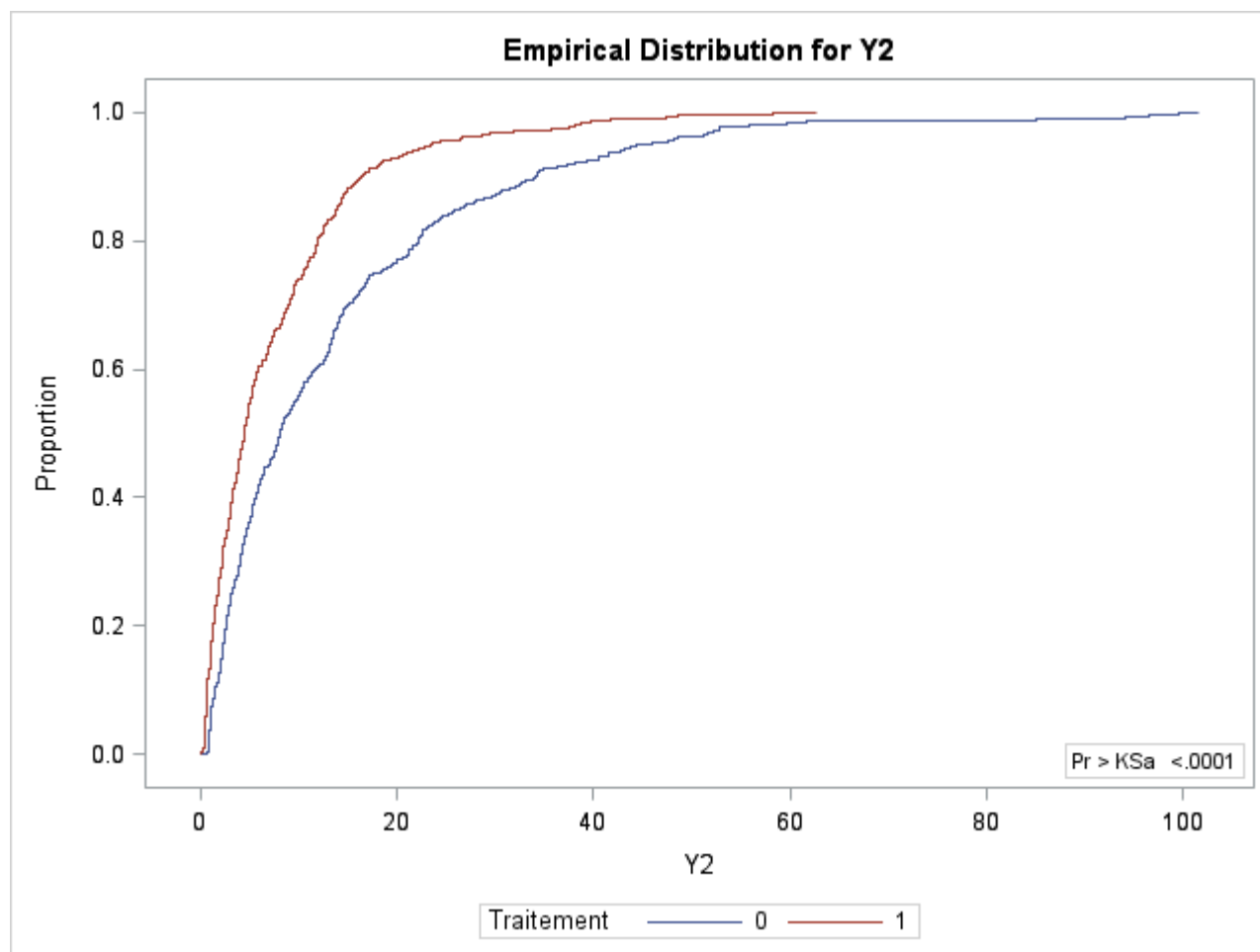
$$W_1 = \frac{(0.3355743)^2}{(0.6426280)^2} = 0,27 < \chi_{(1)}^2 = 3.84 \rightarrow \text{on ne rejette pas } H_0^{(1)}.$$

$$W_2 = \frac{(-7.0531795)^2}{(4.8157434)^2} = 2,15 < \chi_{(1)}^2 = 3.84 \rightarrow \text{on ne rejette pas } H_0^{(2)}.$$

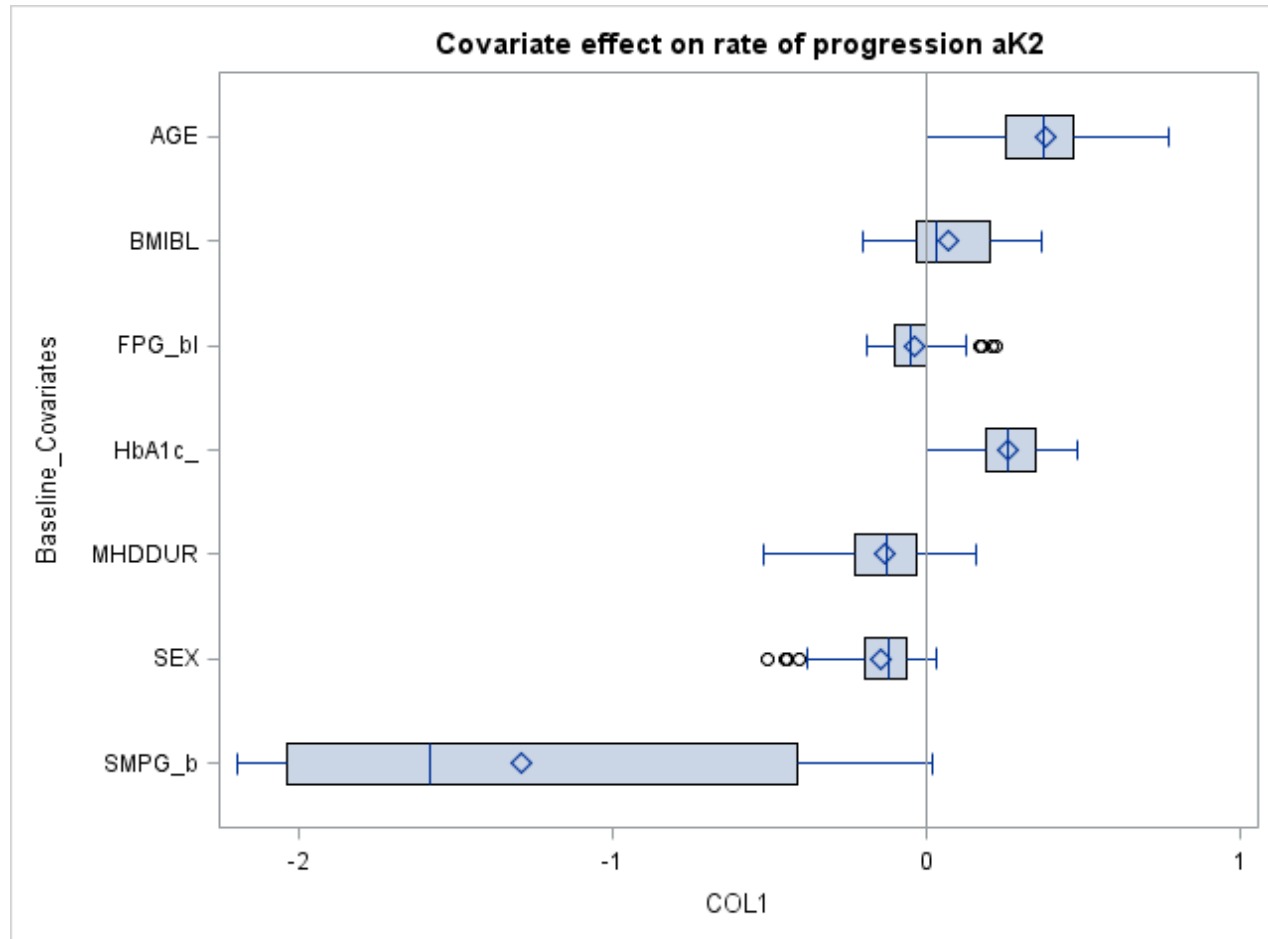
MODELISATION DU DIABETE



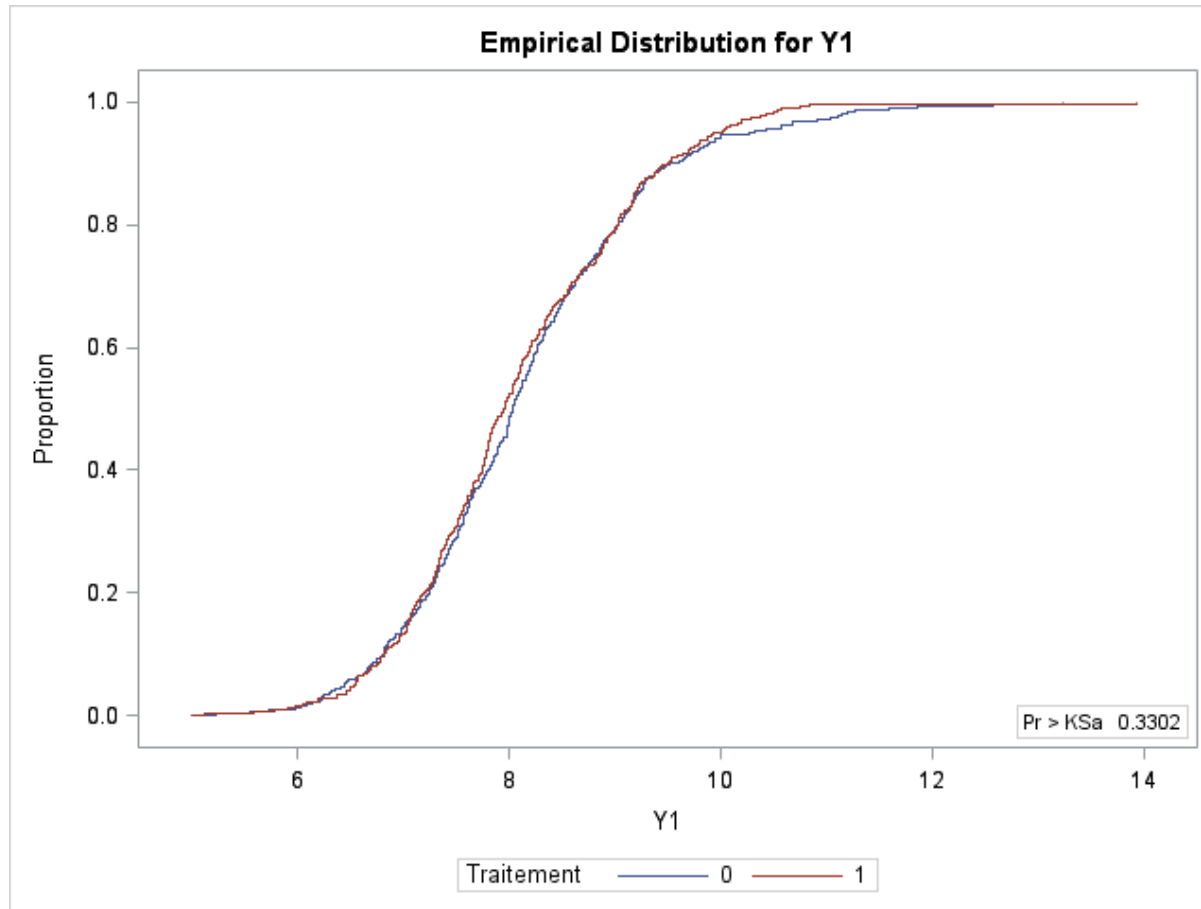
MODELISATION DU DIABETE



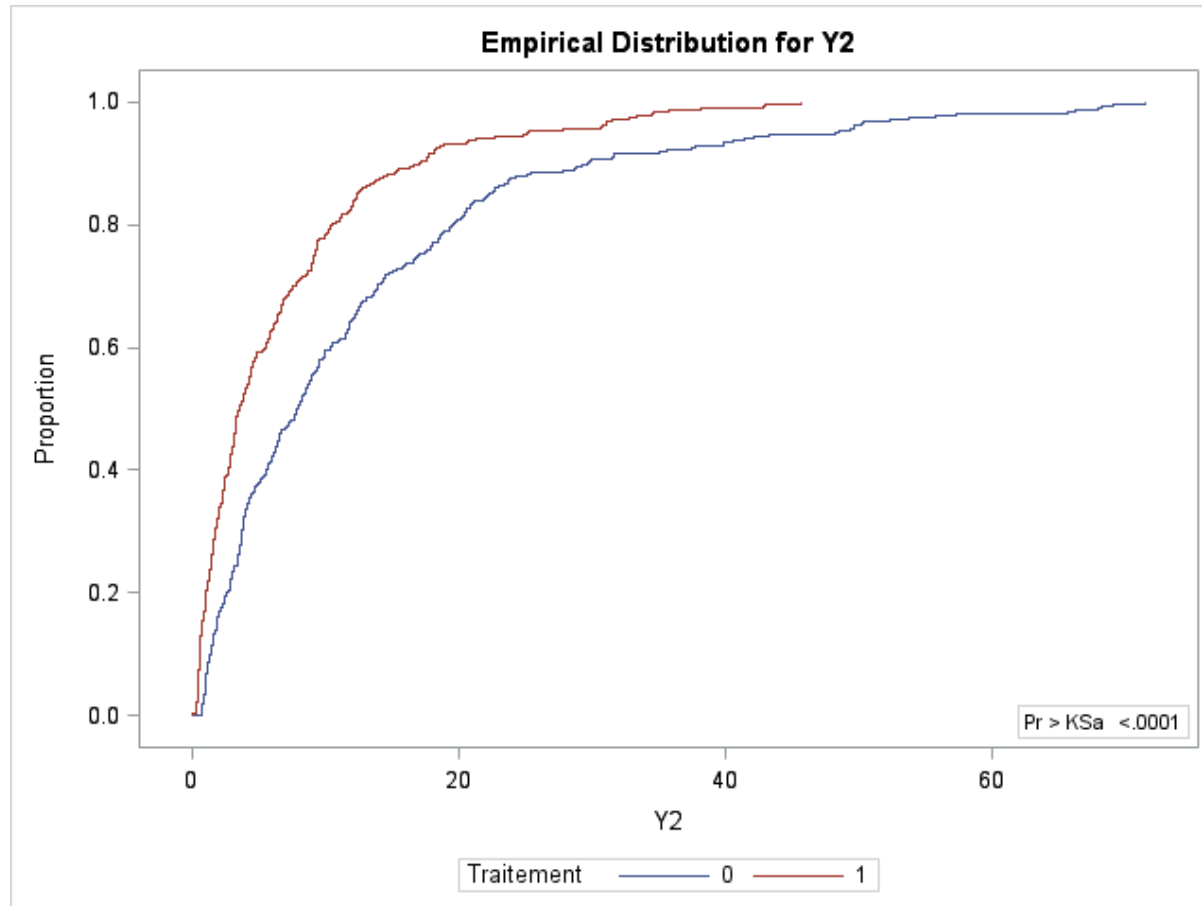
DIABETES MODELING: 24-hour mean SMPG (at month 6) and Hypoglycemia (54 mg/dL severe or confirmed during 26-week on treatment period)



DIABETES MODELING: 24-hour mean SMPG (at month 6) and Hypoglycemia (54 mg/dL severe or confirmed during 26-week on treatment period)



DIABETES MODELING: 24-hour mean SMPG (at month 6) and Hypoglycemia (54 mg/dL severe or confirmed during 26-week on treatment period)



DIABETES MODELING: 24-hour mean SMPG (at month 6) and Hypoglycemia (54 mg/dL severe or confirmed during 26-week on treatment period)

Variable	N	Mean	Std Dev	Minimum	Maximum
beta1	1800	-0.0882438	0.2309073	-0.4604824	1.3022328
beta2	1800	-9.2306856	6.0163732	-18.4944498	0.1469601

$$H_0^{(1)} : \mu_{\beta_1} = 0$$

$$H_0^{(2)} : \mu_{\beta_2} = 0$$

By applying Wald test on each effect, we obtain:

$$W_1 = \frac{(-0.0882438)^2}{(0.2309073)^2} = 0.15 < \chi_{(1)}^2 = 3.84 \rightarrow \text{we do not reject } H_0^{(1)}.$$

$$W_2 = \frac{(-9.2306856)^2}{(6.0163732)^2} = 2.35 < \chi_{(1)}^2 = 3.84 \rightarrow \text{we do not reject } H_0^{(2)}.$$

DIABETES MODELING: 24-hour mean SMPG (at month 6) and Hypoglycemia (54 mg/dL severe or confirmed during 26-week on treatment period)

Traitement=0

Variable	N	Mean	Std Dev	Minimum	Maximum
Y1	501	8.1426031	1.1688611	4.9897156	13.2218719
Y2	501	12.5500409	13.9603706	0.6526336	71.4913199

Traitement=1

Variable	N	Mean	Std Dev	Minimum	Maximum
Y1	475	8.0750205	1.0741337	5.1154162	13.9318057
Y2	475	6.7740773	8.1954945	0.0348566	45.6324341