



HAL
open science

Recherche d'effets des variables liées aux apnées du sommeil sur les facteurs endothéliaux. Influence des traitements

Julie Barbe

► **To cite this version:**

Julie Barbe. Recherche d'effets des variables liées aux apnées du sommeil sur les facteurs endothéliaux. Influence des traitements. Méthodologie [stat.ME]. 2010. dumas-00517789

HAL Id: dumas-00517789

<https://dumas.ccsd.cnrs.fr/dumas-00517789>

Submitted on 15 Sep 2010

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.



BARBE Julie

Master 1 de Mathématiques

Mention Statistique et applications

Université de Strasbourg

**Recherche d'effets des variables liées aux apnées
du sommeil sur les facteurs endothéliaux
Influence des traitements**

Responsable de diplôme : Mme Armelle Guillou

Maître de stage : M. André Malan

Sommaire

Remerciements.....	4
Résumé.....	5
1 Introduction	6
1.1 Présentation du laboratoire	6
1.1.1 L'INCI.....	6
1.1.2 Le département de neurobiologie des rythmes	6
1.2 Présentation du sujet	7
1.2.1 Les troubles respiratoires obstructifs du sommeil.....	7
1.2.2 Présentation des données.....	8
1.2.3 Objectif	9
2 Effets des prédicteurs sur les variables dépendantes	11
2.1 Régression linéaire multiple	11
2.1.1 Analyse détaillée pour la variable ICAM(1).....	11
2.1.2 Résultats de l'analyse des autres facteurs endothéliaux.....	16
2.1.3 Colinéarité des régresseurs	18
2.2 Analyse en composantes principales.....	19
2.2.1 Principe.....	19
2.2.2 Application	20
2.3 Régression linéaire multiple avec nos nouvelles variables	22
3 Différences entre le début et la fin de la nuit.....	25
3.1 Analyse détaillée pour le variable ICAM.....	25
3.1.1 Modèle	25
3.1.2 Vérification des hypothèses	26
3.1.3 Résultats	28
3.2 Résultats des analyses pour les autre facteurs endothéliaux	29
4 Effets des traitements.....	31
4.1 Traitement par PPC.....	31
4.1.1 Effet d'une nuit de traitement par PPC.....	31
4.1.2 Effet d'un mois de traitement PPC.....	34

4.2	Traitement par O2	36
5	Conclusion.....	42
6	Annexe_1	43
7	Annexe_2	44
8	Sources.....	45

Remerciements

Je tiens à remercier Mme Valérie SIMONNEAUX, responsable du département de Neurobiologie des rythmes, de m'avoir permis de réaliser mon stage au sein du laboratoire.

Je souhaite également remercier M. André MALAN, mon maître de stage, pour sa disponibilité et ses conseils à chaque difficulté rencontrée.

Merci également au Pr Patrice BOURGIN et au Dr Elisabeth RUPPERT de m'avoir permis de réaliser cette étude en collaboration avec eux.

Résumé

J'ai effectué mon stage à l'Institut des Neurosciences Cellulaires et Intégratives (INCI), au sein du laboratoire de Neurobiologie des Rythmes.

Durant ce stage, j'ai effectué une étude sur les troubles respiratoires du sommeil et leurs impacts sur les protéines du sang.

Le principal objectif a été d'observer les effets des variables du sommeil et conséquences des troubles du sommeil sur les différents facteurs endothéliaux, puis de regarder les effets sur ceux-ci d'une nuit de sommeil, et enfin d'observer l'effet des traitements utilisés pour soigner les troubles du sommeil sur les variables du sommeil ainsi que leurs répercussions sur les facteurs endothéliaux.

Afin de réaliser cette étude, j'ai utilisé différentes méthodes vues en cours comme la régression linéaire multiple et l'analyse de la variance, ainsi que l'analyse en composantes principales.

1 Introduction

1.1 Présentation du laboratoire

1.1.1 L'INCI

L'Institut des Neurosciences Cellulaires et Intégratives (INCI) est un laboratoire du CNRS associé à l'Université de Strasbourg, dont les études visent à comprendre le fonctionnement des cellules nerveuses et neuroendocrines (qui libèrent une substance hormonale dans le système nerveux), et des circuits neuronaux.

Les études sont menées suivant une approche multidisciplinaire et caractérisée par différents niveaux d'investigation : génomique, protéomique, cellulaire, intégré et comportemental.

L'institut se compose de douze équipes de recherche réparties dans trois départements :

- Neurobiologie des rythmes
- Neurotransmission et sécrétion neuroendocrine
- Nociception et douleur

1.1.2 Le département de neurobiologie des rythmes

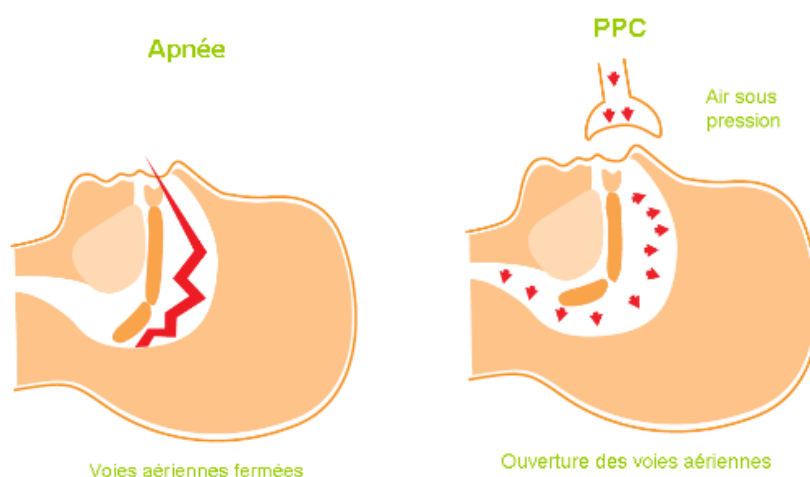
Les recherches conduites par ce département ont pour objectif de comprendre les mécanismes nerveux et neuroendocrines impliqués dans le contrôle des rythmes biologiques. Ces rythmes permettent à l'organisme de s'adapter aux variations journalières et saisonnières de l'environnement. Tout dysfonctionnement causé par-ou associé à- une déstructuration des rythmes biologiques entraîne des troubles graves que l'on trouve, par exemple, chez les travailleurs postés, les personnes souffrant de dépressions saisonnières, des insomniaques, des aveugles, ou bien encore chez les personnes âgées. Au laboratoire, cette thématique est abordée chez différentes espèces de mammifères, dont le hamster et le rat. Une des équipes de ce département étudie le sommeil en association avec le service du sommeil des hôpitaux universitaires. C'est là qu'ont été recueillies les données du présent travail.

1.2 Présentation du sujet

1.2.1 Les troubles respiratoires obstructifs du sommeil

Ces troubles qui comprennent le syndrome des apnées hypopnées obstructives du sommeil correspondent à un spectre d'affections respiratoires nocturnes qui se caractérisent essentiellement par l'affaissement partiel ou complet du pharynx (carrefour entre les voies aériennes et les voies digestives) au moment de l'endormissement. L'affaissement complet correspond à l'apnée qui interdit tout passage de l'air de 10 secondes à plus d'1 minute et peut se répéter plusieurs fois au cours d'une heure de sommeil. Les apnées peuvent être précédées ou suivies par des affaissements partiels, on parle alors d'hypopnées. Les conséquences biologiques de ces apnées sont une diminution du taux d'oxygène dans le sang et l'apparition d'une somnolence diurne excessive due à un fractionnement du sommeil par de multiples microéveils (la plupart du temps invisibles). De plus, ces troubles respiratoires du sommeil constituent un facteur de risque de complications cardiovasculaires, ils pourraient en particulier altérer la fonction de l'endothélium (couche la plus interne des vaisseaux sanguins, en contact avec le sang) vasculaire.

Le traitement le plus utilisé pour soigner le syndrome d'apnées du sommeil est la pression positive continue (PPC). Cette assistance nocturne consiste en un envoi d'air à une pression suffisante pour maintenir les voies aériennes supérieures ouvertes malgré la diminution du tonus musculaire dilatateur du pharynx dans le sommeil. Elle permet alors, en corrigeant à la fois l'obstruction pharyngée et les épisodes d'hypoxie, de rétablir une respiration nocturne normale ainsi qu'un sommeil de bonne qualité.



Un deuxième traitement sera vu lors de cette étude : le traitement par oxygénation nasale. Il permet de corriger les épisodes d'hypoxie mais pas l'obstruction pharyngée toujours responsable de la fragmentation du sommeil contrairement au traitement par PPC.

1.2.2 Présentation des données

Cette étude a été réalisée en collaboration avec le Pr Bourgin et le Dr Ruppert du service de neurologie de l'Hôpital Civil de Strasbourg.

Nous disposons d'observations recueillies sur 43 patients présentant une gamme étendue de troubles liés aux apnées du sommeil. Nous avons trois types de variables différentes : des variables générales caractérisant l'état des patients, des variables décrivant le sommeil et les conséquences des apnées sur l'oxygénation du sang enregistrées au cours d'une nuit de sommeil à la clinique, et des facteurs endothéliaux (nos variables dépendantes) mesurés avant et après la nuit d'enregistrement.

Variables générales :

- Age : variable quantitative continue,
- Sexe : variable qualitative,
- Indice de masse corporelle *BMI* : variable quantitative continue,
- Score de somnolence Epworth *EPS* (questionnaire qui mesure la somnolence diurne) : variable quantitative continue,
- Hypertension artérielle *HTA* : variable quantitative continue,
- Traitement de l'hypertension artérielle *TTT* : variable qualitative,
- β -bloquants *beta_c* : variable qualitative,
- Inhibiteurs *All inh_all* : variable qualitative,
- Hypercholestérolémie *Hchol_c* : variable qualitative,
- Hypertriglycéridémie *Trigl_c* : variable qualitative,
- Tabac *tabac_c* : variable qualitative.

Variables du sommeil et conséquences des troubles du sommeil:

- Indice d'apnée hypopnée *IAH* : variable quantitative continue,
- Nombre total d'évènements *n_tot_evt* : variable quantitative discrète,
- Index de limitation de débit respiratoire *ILD_h* : variable quantitative continue,
- Diagnostic *diag_c4* : variable qualitative,
- Temps total de sommeil *TTS* (en min) : variable quantitative continue,
- Effet sommeil *Eff_Som%* (=temps sommeil/temps au lit) : variable quantitative continue,

- Pourcentage de ronflements *R%* : variable quantitative continue,
- Pourcentage de sommeil lent profond *SLP%* : variable quantitative continue,
- Pourcentage de sommeil paradoxal *SP%* : variable quantitative continue,
- Veille *V_min* (en min) : variable quantitative continue,
- Temps de saturation < 90% *tdesat%* : variable quantitative continue,
- Saturation moyenne *sat_moy%* : variable quantitative continue,
- Nombre total de désaturations *n_tot_desat* : variable quantitative discrète,
- Désaturations par heure de sommeil *n_desat_h* : variable quantitative continue,
- Type de traitement *traitmt* : variable qualitative.

A ce stade de l'étude, nous n'avons pas de raison de faire une distinction entre ces variables pour répondre aux questions posées.

Facteurs endothéliaux (soir(1) et matin(2)) :

- Tissu plasminogène *t-PA* : variable quantitative continue,
- Inhibiteur de l'activateur plasminogène *PAI* : variable quantitative continue,
- Thrombomoduline *TBM* : variable quantitative continue,
- Adhésion cellulaire *ICAM* : variable quantitative continue,
- Adhésion vasculaire *VCAM* : variable quantitative continue,
- Sélectine endothéliale *e-SEL* : variable quantitative continue,
- Sélectine plaquettaire *p-SEL* : variable quantitative continue,
- Interleukin 18 *IL_18* : variable quantitative continue,
- Facteur de von Willebrand *VWF* : variable quantitative continue,
- Fibrinogène *FG* : variable quantitative continue.

Un sous-groupe de patients a été suivi à différentes reprises après avoir subi un ou plusieurs traitements (PPC et/ou oxygénation nasale) destinés à corriger leur état.

1.2.3 Objectif

Afin de répondre aux questions posées, nous allons dans un premier temps nous intéresser aux effets des prédicteurs (variables générales et du sommeil) sur les variables dépendantes (facteurs endothéliaux), puis nous nous attarderons sur les différences entre le début et la fin de la nuit de sommeil, et enfin nous nous demanderons si les traitements utilisés ont un

effet sur les différents aspects du sommeil et si ceux-ci se répercutent sur les variables dépendantes identifiées précédemment.

L'ensemble de l'étude sera réalisée à l'aide du logiciel Statistica.

2 Effets des prédicteurs sur les variables dépendantes

Dans le but de répondre à cette première question, nous effectuerons tout d'abord des régressions linéaires multiples pour chaque variable dépendante, puis, afin de préciser et améliorer les résultats trouvés, nous utiliserons une analyse en composantes principales (ACP).

2.1 Régression linéaire multiple

2.1.1 Analyse détaillée pour la variable ICAM(1)

On s'intéresse au modèle de régression multiple suivant :

$$Y_i = \beta_0 + \sum_{j=1}^{26} \beta_j x_{ij} + \varepsilon_{ij}, 1 \leq i \leq 33.$$

Où Y_i représente la variable *ICAM(1)* (adhésion cellulaire, mesurée le soir pour le patient i) et les x_{ij} représentent respectivement, pour le patient i , les variables générales *age*, *sexe*, *BMI*, *EPS*, *HTA*, *TTT*, *beta_c*, *inh_all*, *Hchol_c*, *Trigl_c*, *tabac_c* et les variables du sommeil et conséquences des troubles du sommeil *IAH*, *n_tot_evt*, *ILD_h*, *diag_c4*, *TTS*, *Eff_som%*, *R%*, *SLP%*, *SP%*, *V_min*, *tdesat%*, *sat_moy%*, *n_tot_desat*, *n_desat_h* et *traitmt*.

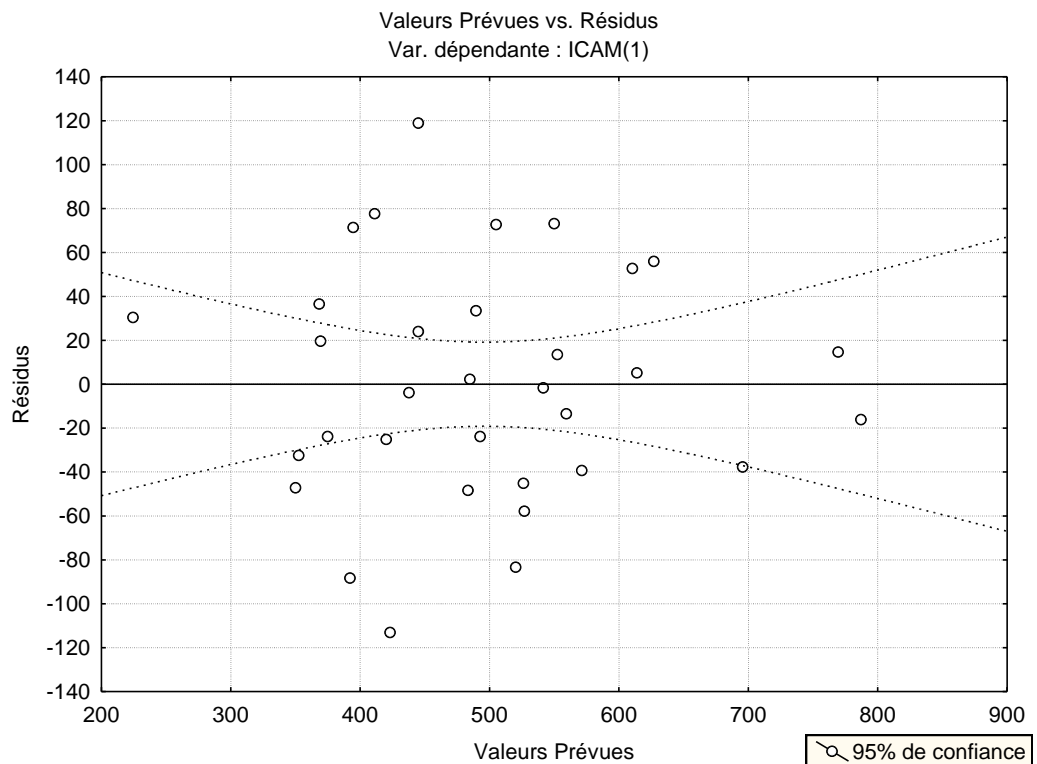
On remarquera que le nombre d'observations ici est de 33 au lieu de 43, ce qui s'explique par le fait que le logiciel Statistica ne prend pas en compte les observations ayant certaines valeurs manquantes.

Les hypothèses suivantes doivent être vérifiées : les erreurs doivent être indépendantes, centrées, de même variance (condition d'homoscédasticité) et doivent suivre une loi normale.

On commence par vérifier les hypothèses :

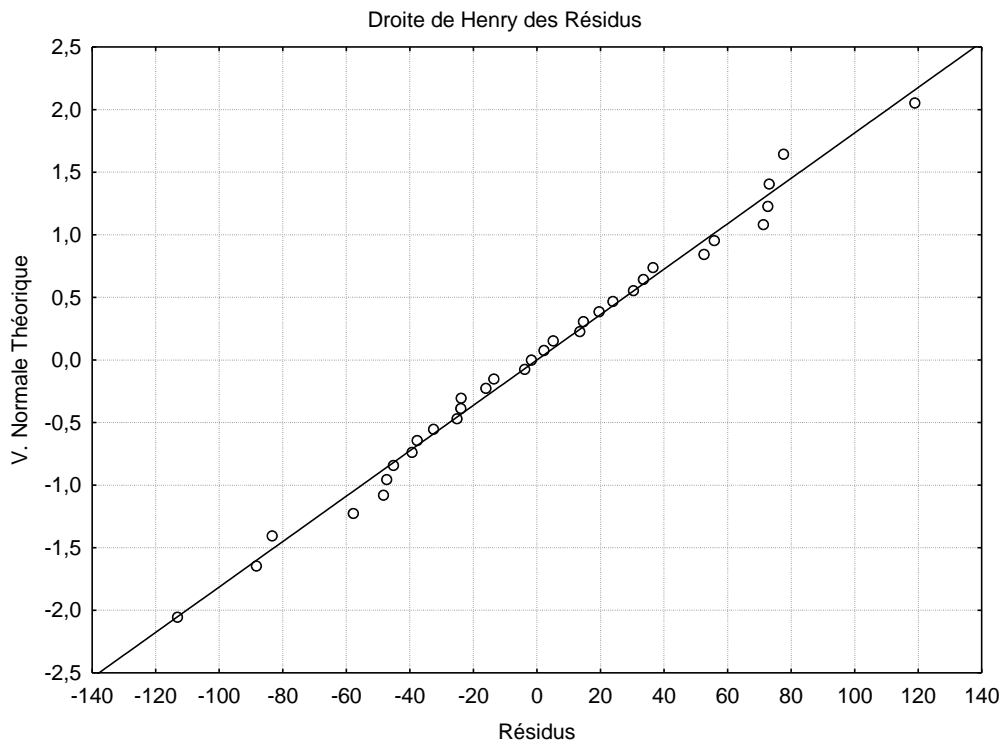
- D'après le protocole expérimental, on peut supposer que les patients sont tous différents et ne sont pas liés, l'hypothèse d'indépendance des erreurs est donc vérifiée.
- On a inclus une constante β_0 dans notre modèle, on peut donc en déduire que les erreurs sont centrées.
- On ne peut pas se servir du test de Bartlett pour tester l'hypothèse d'homogénéité des variances car on n'a qu'une seule observation pour chaque patient (pas de

mesures répétées). On va donc vérifier graphiquement cette hypothèse : on trace le nuage de points des résidus en fonction des valeurs prévues.



On remarque que les résidus sont répartis dans une bande horizontale autour de 0, on peut alors en conclure que l'hypothèse d'homogénéité des variances est bien vérifiée.

- Pour vérifier la normalité des erreurs, on va se servir de la droite de Henry. La droite de Henry est une méthode graphique qui ajuste une distribution gaussienne à celle d'une série d'observations (d'une variable numérique continue).



On remarque que les résidus sont bien alignés sur la droite, on peut en déduire que l'hypothèse de normalité est bien vérifiée.

Toutes les hypothèses du modèle sont bien vérifiées, on regarde maintenant si le modèle est intéressant.

On va donc tester $H_0 : \beta_1 = \dots = \beta_{26} = 0$ contre l'hypothèse alternative.

Stats de Synthèse; VD: ICAM(1) (SAS + COR bio version 8_4_10_v4)	
	Valeur
R Multiple	0,9166
R² Multiple	0,8401
R² Ajusté	0,1473
F(26,6)	1,2126
p	0,4387
Err-Type de l'Estim.	122,6753

On trouve la p-valeur associée, par la statistique de Fisher, à l'hypothèse nulle H_0 égale à 0.4387 >> 0.05 donc on ne peut pas rejeter H_0 au seuil 5%.

On s'intéresse maintenant aux effets de chacun des facteurs explicatifs. On teste $\forall 1 \leq j \leq 26$, $H_0 : \beta_j = 0$ contre $H_1 : \beta_j \neq 0$. Toutes les p-valeurs associées aux hypothèses nulles sont strictement supérieures à 0.05, on ne peut donc pas rejeter H_0 au seuil 5%. Aucune des variables explicatives du modèle n'est significative. Le modèle n'est donc pas un bon modèle.

Afin de trouver un modèle plus adapté, on va utiliser une méthode de choix de modèle descendante. Cela nous donne le modèle suivant :

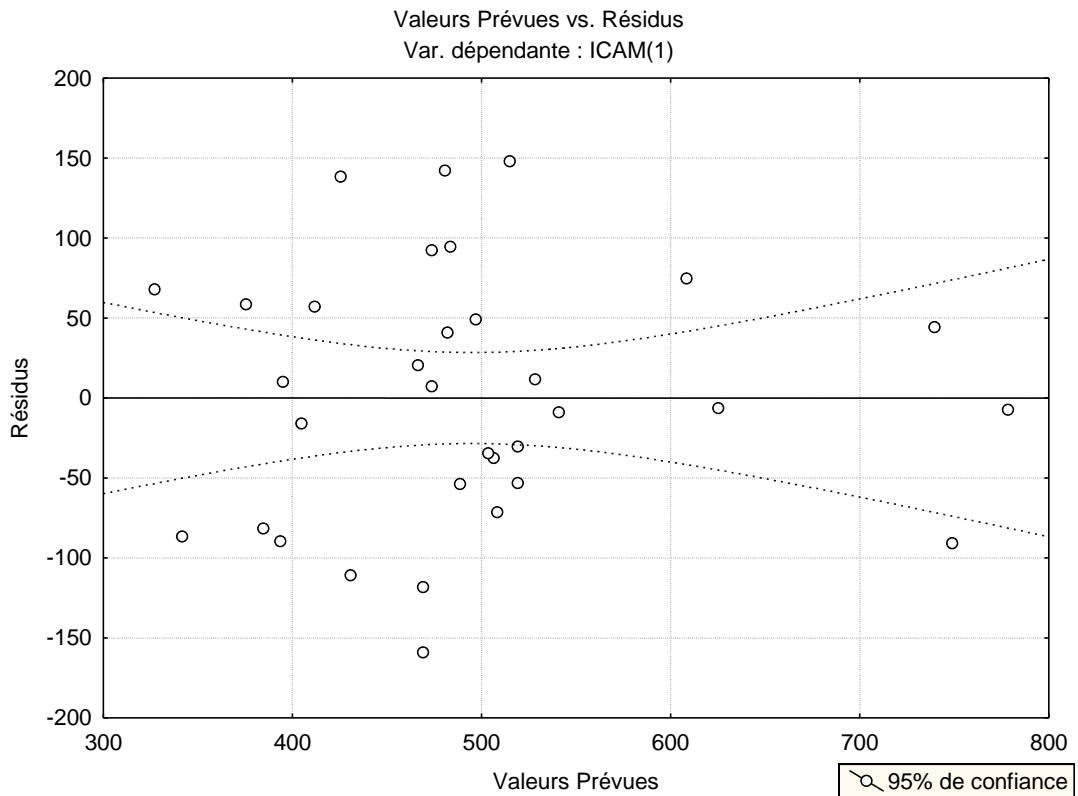
$$Y_i = \beta_0 + \sum_{j=1}^6 \beta_j x_{ij} + \varepsilon_{ij}, 1 \leq i \leq 33.$$

Où Y_i représente toujours la variable *ICAM(1)* (adhésion cellulaire, mesurée le soir pour le patient i) et les x_{ij} représentent respectivement, pour le patient i, les variables générales *Hchol_c* et *tabac_c* et les variables du sommeil et conséquences des troubles du sommeil *n_tot_evt*, *R%*, *SLP%* et *n_desat_h*.

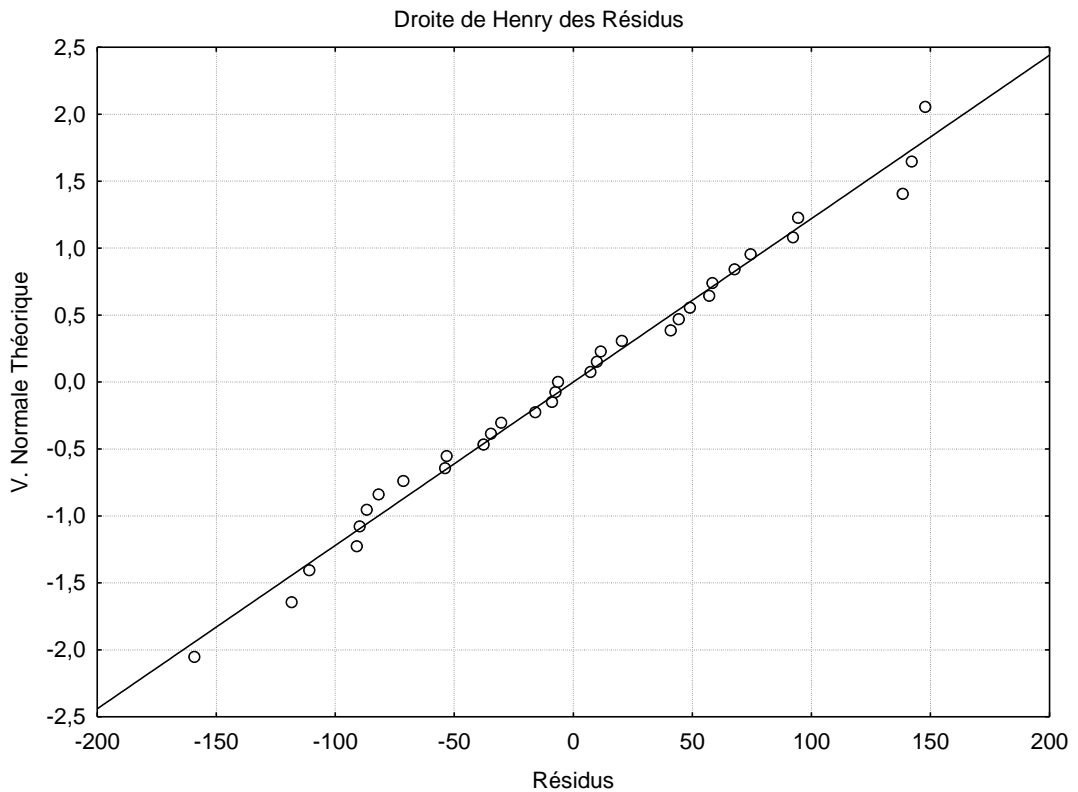
On vérifie les hypothèses de ce nouveau modèle :

- L'hypothèse d'indépendance des erreurs est toujours vérifiée car on a le même protocole expérimental.
- On a inclus une constante β_0 dans le modèle donc on peut supposer que les erreurs sont centrées.

- On vérifie graphiquement l'hypothèse d'homogénéité des variances en traçant le nuage de points des résidus en fonction des valeurs prévues :



- On utilise la droite de Henry pour vérifier l'hypothèse de normalité des erreurs :



Les résidus sont bien alignés sur la droite, on peut en déduire que l'hypothèse de normalité est bien vérifiée.

Toutes les hypothèses du modèle sont donc vérifiées. On va maintenant tester l'hypothèse $H_0 : \beta_1 = \dots = \beta_6 = 0$ contre l'hypothèse alternative.

Stats de Synthèse; VD: ICAM(1) (SAS + COR bio version 8_4_10_v4)	
	Valeur
R Multiple	0,80463
R² Multiple	0,64742
R² Ajusté	0,56606
F(6,26)	7,95711
p	0,00006
Err-Type de l'Estim.	87,51394

On trouve la p-valeur associée, par la statistique de Fisher, à l'hypothèse nulle H_0 égale à $0.00006 \ll 0.05$ donc on ne peut pas accepter H_0 au seuil 5% (le risque associé cette décision est un risque de première espèce $\alpha = 0.05$) et $R^2 = 64.74\%$ ce qui est très bon. Le modèle est donc bien explicatif. De plus toutes les variables explicatives du modèle sont significatives au niveau 0.05.

On a répété le même schéma pour tous les facteurs endothéliaux, on ne présentera pas le détail dans ce rapport par souci de présentation et de clarté mais on donnera les principaux résultats.

2.1.2 Résultats de l'analyse des autres facteurs endothéliaux

Pour chaque variable dépendante, on donnera les variables explicatives du modèle trouvé par la méthode de choix de modèle ascendante, avec la p-valeur associée au test global $H_0 : \beta_1 = \dots = \beta_p = 0$ contre l'hypothèse alternative, et le R^2 de la régression respective. On précisera également les variables significatives du modèle trouvées grâce aux tests individuels $H_0 : \beta_j = 0$ contre l'hypothèse alternative.

Variable dépendante	Variables explicatives	p-valeur	R²	Variables significatives (au seuil 5%)
t-PA(1)	<i>Sexe, n_tot_evt, TTS, Eff_Som%, R%, SLP%, V_min, n_desat_h</i>	0.00004	70.80%	toutes
t-PA(2)	<i>Age, BMI, tabac_c, Trigl_c, TTT, n_tot_evt, SLP%, traitmt, n_desat_h</i>	0.00002	65%	<i>BMI, Trigl_c, tabac_c, n_desat_h</i>
PAI(1)	Aucun modèle approprié trouvé			
PAI(2)	<i>BMI, tabac_c, EPS, Trig_c, IAH, R%, SP%, n_tot_desat, tdesat%, sat_moy%, traitmt</i>	0.00000	83.32%	<i>BMI, Tabac_c, EPS, IAH, SP%, Sat_moy%, n_tot_desat</i>
TBM(1)	Aucun modèle approprié trouvé			
TBM(2)	Aucun modèle approprié trouvé			
ICAM(2)	<i>Tabac_c, inh_all, R%, SLP%</i>	0.00002	53.3%	<i>Tabac_c, inh_all</i>
VCAM(1)	<i>BMI, tabac_c, EPS, TTT, Trigl_c, HTA, ILD_h, inh_all, n_desat_h, traitmt</i>	0.0011	66.92%	<i>BMI, EPS, TTT, ILD_h</i>
VCAM(2)	<i>Sexe, BMI, Hchol_c, TTT, IAH, IDL_h, inh_all, n_tot_evt, tdesat%, sat_moy%, traitmt</i>	0.0025	68.54%	Sexe, BMI
e-SEL(1)	<i>BMI, sat_moy%</i>	0.00000	72.73%	<i>BMI, sat_moy%</i>
e-SEL(2)	<i>BMI, n_tot_desat</i>	0.00000	68.24%	<i>BMI, n_tot_desat</i>
p-SEL(1)	<i>Tabac_c, Trigl_c, IAH, SLP%, tdesat%</i>	0.04807	33.47%	<i>tdesat%</i>
p-SEL(2)	Aucun modèle approprié trouvé			
IL18(1)	Aucun modèle approprié trouvé			
IL18(2)	Aucun modèle approprié trouvé			
VWFT(1)	<i>Sexe, BMI, EPS, inh_all, Hchol_c, Trigl_c, tabac_c, ILD_h, R%, SP%, n_desat_h, traitmt</i>	0.00208	78.45%	<i>BMI, ILD_h, inh_all, n_desat_h</i>
VWFT(2)	<i>Sexe, BMI, EPS, inh_all, Trigl_c, HTA, IAH, n_tot_evt, ILD_h, SLP%, SP%, traitmt</i>	0.00048	79.41%	<i>BMI, HTA</i>
FGT(2)	Aucun modèle approprié trouvé			

Pour certaines variables dépendantes, on n'a pas trouvé de modèle explicatif approprié. Cela vient soit d'une mauvaise vérification des hypothèses, soit d'un manque de données. Dans certaines régressions, on a pu remarquer la présence de quelques valeurs atypiques (données directement par Statistica), qui peuvent venir d'imprécisions de machines lors des mesures par exemple. Selon les cas, celles-ci ont été gardées ou retirées de la régression.

Pour chaque modèle, les hypothèses sont bien vérifiées (on a procédé de la même manière que pour l'étude de la variable *ICAM(1)*) et on a très peu de cas d'observations atypiques. Afin de valider nos modèles, il reste à vérifier la colinéarité des régresseurs.

2.1.3 Colinéarité des régresseurs

Afin de voir si un problème de colinéarité existe dans nos modèles, on va utiliser la « Variance Inflation Factor » (VIF). Pour chaque variable X_j , on a $VIF_j = \frac{1}{1 - R_j^2}$ où R_j^2 est le coefficient de corrélation multiple de la régression avec constante sur toutes les autres variables prédictives. S'il y a colinéarité, R_j^2 est proche de 1, donc VIF_j est grand. D'après la règle de Besley et al., une valeur de VIF plus grande que 10 révèle un problème. Afin de trouver les VIFs, Statistica nous donne la « Matrice sweep courante » (annexe_1) et à partir de là, les VIFs sont l'opposé des éléments diagonaux de cette matrice. Ici, on a trouvé les valeurs de VIF suivantes:

Variables	Valeurs de VIF
<i>age</i>	7.8909
<i>sexe</i>	4.54602
<i>BMI</i>	6.24703
<i>EPS</i>	5.3544
<i>HTA</i>	23.0111
<i>TTT</i>	7.6499
<i>beta_c</i>	5.29292
<i>Inh_all</i>	17.5677
<i>Hchol_c</i>	17.0056
<i>Trigl_c</i>	2.02072
<i>tabac_c</i>	4.7276
<i>IAH</i>	442.186
<i>n_tot_evt</i>	287.274
<i>ILD_h</i>	59.224
<i>diag_c4</i>	34.2957
<i>TTS</i>	47.6970
<i>Eff_som%</i>	445.225

R%	8.0811
SLP%	2.3802
SP%	11.5043
V_min	372.512
tdesat%	23.5719
sat_moy%	35.7869
n_tot_desat	150.215
n_desat_h	175.219
traitmt	3.55693

On remarque que 15 de nos variables ont une valeur de VIF grande, certaines excédant largement 10. On a donc un problème de multi-colinéarité dans nos modèles, ce qui est logique car certaines redondances dans les variables sont évidentes, comme par exemple les variables TTS et Eff_som% (Eff_som%=TTS/temps au lit).

Pour améliorer les résultats trouvés, on va utiliser une analyse en composantes principales (ACP).

2.2 Analyse en composantes principales

2.2.1 Principe

Les analyses factorielles forment une famille de méthodes pour explorer des données multidimensionnelles et essayer de réduire la dimension du problème. L'analyse en composantes principales (ACP) est un type d'analyse factorielle. On dispose de n observations décrites par p variables numériques x_1, \dots, x_p et on souhaite réduire le nombre de facteurs considérés, tout en conservant le maximum d'information. Ces nouvelles variables seront des combinaisons linéaires des variables d'origine et porteront le nom de composantes principales. Les composantes principales définissent des directions de l'espace des observations qui sont deux à deux orthogonales. Autrement dit, l'ACP procède à un changement de repère orthogonal, les directions originales étant remplacées par les composantes principales.

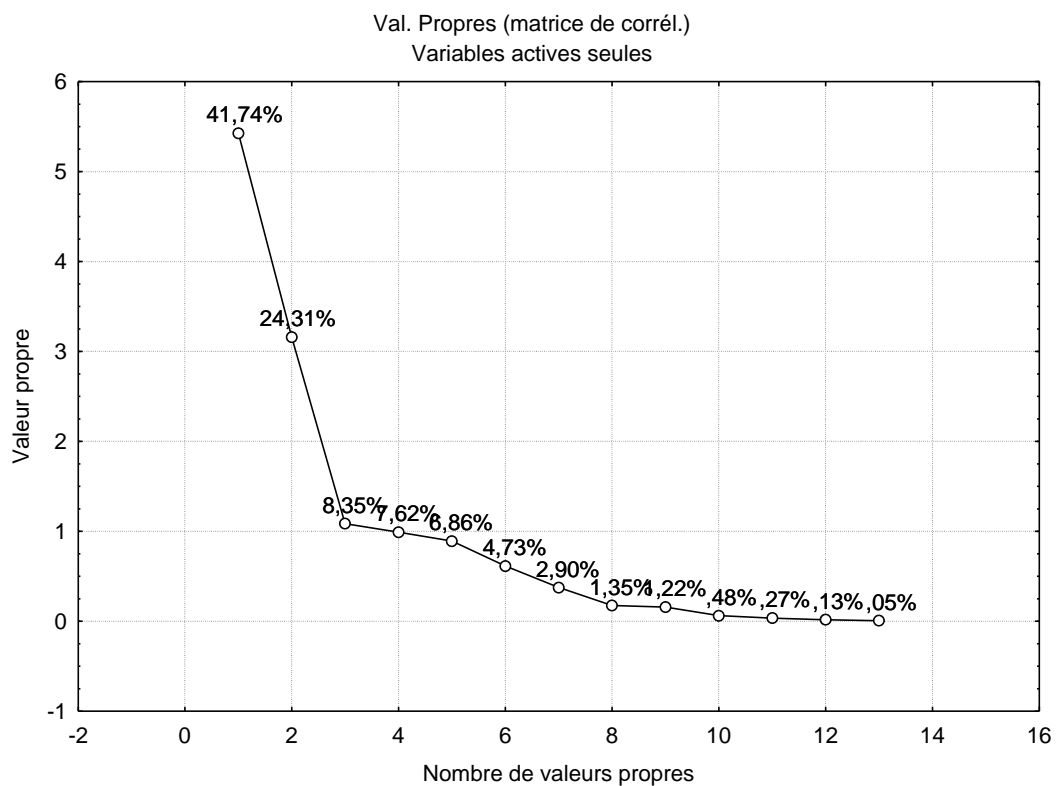
On utilise les données centrées réduites. Les composantes principales sont engendrées par les vecteurs u_1, \dots, u_p , vecteurs propres de la matrice des corrélations R , associés aux valeurs propres $\lambda_1 \geq \dots \geq \lambda_p$. Pour $1 \leq i \leq n$, on a $x_i = (x_{i1}, \dots, x_{ip})$ et le centre de gravité du nuage de points $N = \{x_1, \dots, x_n\}$, $g = (\bar{x}_1, \dots, \bar{x}_p)$. Comme les données sont centrées-réduites, on a $g = (0, \dots, 0)$ et l'inertie totale du nuage de points :

$$I(N, g) = I(N, 0) = \frac{1}{n} \sum_{i=1}^n d^2(x_i, 0) = \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^p x_{ji}^2 = \sum_{j=1}^p \frac{1}{n} \sum_{i=1}^n x_{ji}^2 = \sum_{j=1}^p s_j^2 = \sum_{j=1}^p 1 = p.$$

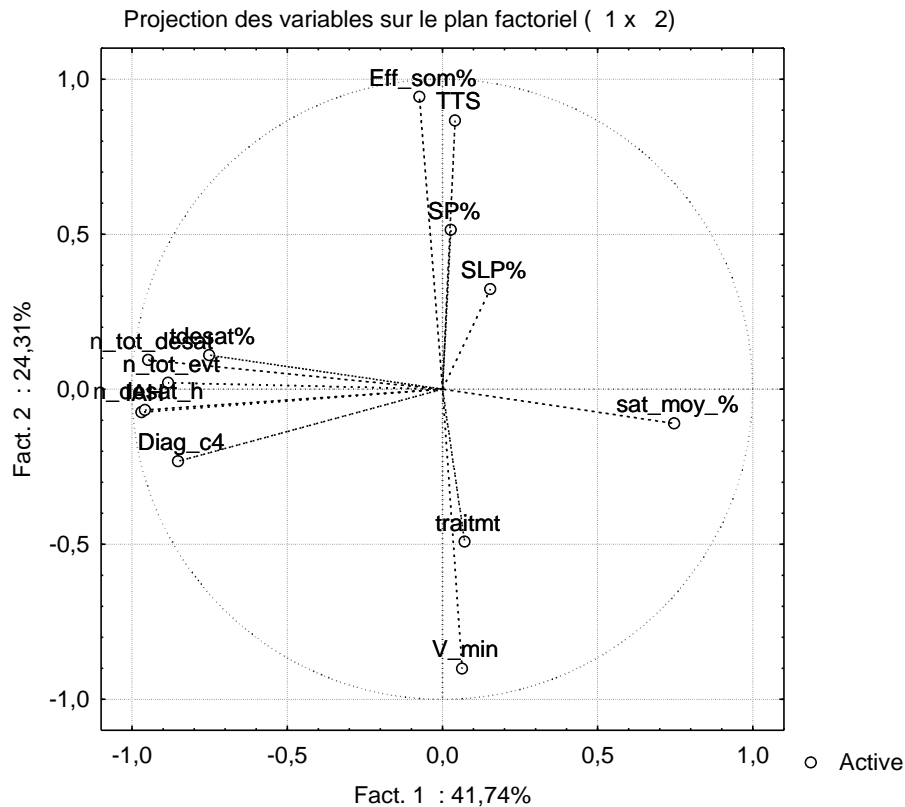
L'inertie du nuage projeté sur l'axe engendré par u_i est égale à λ_i , et la part d'inertie expliquée par l'axe i est égale à $\frac{\lambda_i}{p}$.

2.2.2 Application

On effectue l'ACP sur les variables du sommeil et conséquences des troubles du sommeil. On commence par présenter les valeurs propres en pourcentage d'inertie projetée, afin de choisir le nombre d'axes que nous allons utiliser :



On observe un décrochement (coude) suivi d'une décroissance régulière à partir de la troisième valeur propre (règle du coude : on sélectionne les axes avant le décrochement). De plus, plus de 65% de l'information peut être représentée sur les deux premiers axes principaux, nous allons donc représenter nos variables dans ce plan.

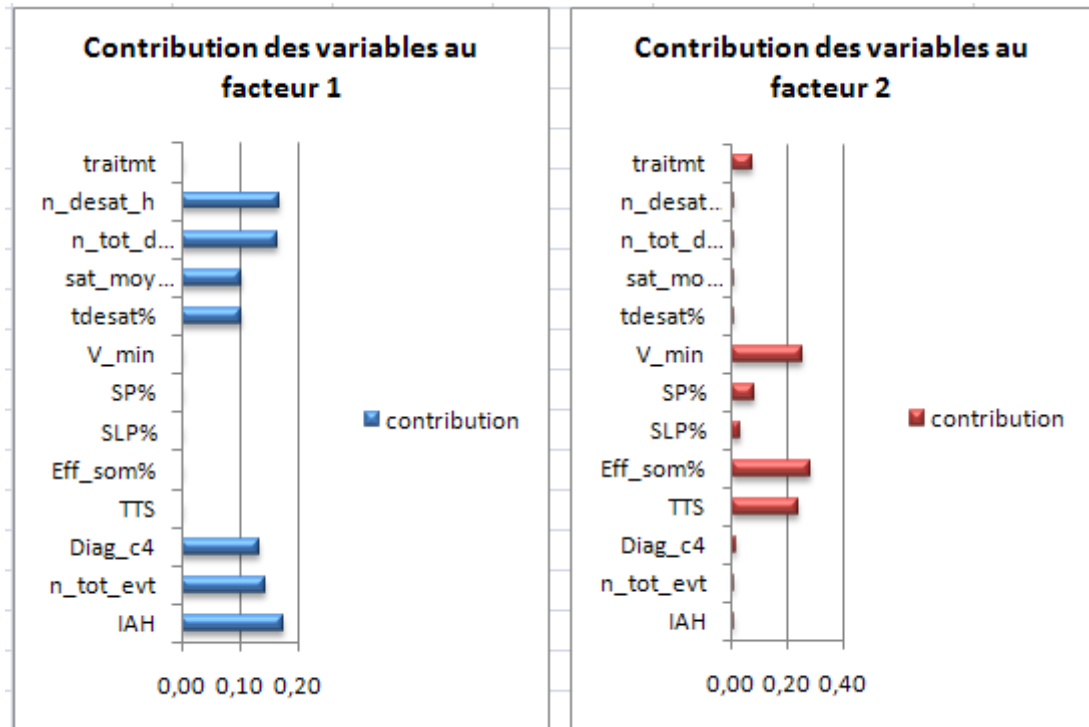


Nous pouvons alors constater que le premier axe est représenté par les variables de saturation du sang en oxygène ainsi que le diagnostic, le nombre d'évènements et l'indice d'apnée hypopnée. On note également que le deuxième axe semble composé des variables du sommeil ainsi que du traitement administré. Cependant, on remarque que la variable *SLP%* n'est pas très bien représentée (angle entre la variable et l'axe plus important), et que les variables *SP%* et *traitmt* ne sont pas très proches de la circonférence du cercle et donc ne contribuent pas très bien à l'inertie globale.

Afin de confirmer cette interprétation graphique, nous allons présenter le tableau des contributions des variables aux deux premiers facteurs, ainsi que les graphiques correspondant :

variables	Contribution au facteur 1	Contribution au facteur 2
<i>IAH</i>	0.173231	0.001734
<i>n_tot_evt</i>	0.144002	0.000140
<i>diag_c4</i>	0.133662	0.017162
<i>TTS</i>	0.000309	0.237806
<i>Eff_som%</i>	0.001000	0.281709
<i>SLP%</i>	0.004407	0.032947
<i>SP%</i>	0.000132	0.083463
<i>V_min</i>	0.000731	0.256680
<i>tdesat%</i>	0.104104	0.003798
<i>sat_moy%</i>	0.102824	0.003850

<i>n_tot_desat</i>	0.165610	0.002806
<i>n_desat_h</i>	0.169045	0.001446
<i>traitmt</i>	0.000944	0.076461



Grâce à ces contributions, on peut créer deux nouvelles variables (*facteur 1* et *facteur 2*) quantitatives continues orthogonales, comme combinaisons linéaires de nos variables initiales. On a ainsi éliminé notre problème de colinéarité entre les variables. Il peut alors être intéressant de reprendre nos régressions linéaires multiples avec ces nouvelles variables prédictrices.

2.3 Régression linéaire multiple avec nos nouvelles variables

On refait donc nos régressions multiples suivant le même schéma que dans la partie 2.1.1.1 et on ne présentera encore ici qu'un tableau des résultats obtenus. Nos variables explicatives sont toujours les variables générales *age*, *sexe*, *BMI*, *EPS*, *HTA*, *TTT*, *beta_c*, *inh_all*, *Hchol_c*, *Trigl_c*, *tabac_c* et nous avons regroupé nos variables du sommeil et conséquences liées aux troubles du sommeil en deux nouvelles variables : *facteur 1* et *facteur 2*. On a trouvé les résultats suivants :

Variable dépendante	Variables explicatives	p-valeur	R²	Variables significatives (au seuil 5%)
t-PA(1)	<i>Age, BMI, tabac_c, HTA, inh_all, facteur 1</i>	0.00006	54.20%	<i>Tabac_c, BMI, HTA, facteur 1</i>
t-PA(2)	<i>Age, BMI, tabac_c, Trigl_c, facteur 1</i>	0.00005	52%	<i>BMI, Trigl_c, facteur 1</i>
PAI(1)	<i>Sexe, EPS, inh_all, HTA, TTT, facteur 2</i>	0.0016	44%	<i>Inh_all, facteur 2</i>
PAI(2)	<i>BMI, tabac_c, inh_all, HTA, facteur 1, facteur 2</i>	0.00000	65.40%	<i>BMI, inh_all</i>
TBM(1)	Aucun modèle approprié trouvé			
TBM(2)	Aucun modèle approprié trouvé			
ICAM(1)	<i>BMI, inh_all, tabac_c, sexe, HTA</i>	0.0044	37.39%	<i>Inh_all, tabac_c</i>
ICAM(2)	Aucun modèle approprié trouvé			
VCAM(1)	<i>Age, BMI, EPS, HTA, TTT, tabac_c</i>	0.0395	30.13%	<i>EPS, BMI</i>
VCAM(2)	Aucun modèle approprié trouvé			
e-SEL(1)	<i>Sexe, BMI, facteur 2</i>	0.00000	68.04%	<i>BMI,</i>
e-SEL(2)	<i>Sexe, age, BMI, Hchol_c, Trigl_c, facteur 1</i>	0.00000	69.70%	<i>BMI, facteur 1</i>
p-SEL(1)	Aucun modèle approprié trouvé			
p-SEL(2)	Aucun modèle approprié trouvé			
IL18(1)	Aucun modèle approprié trouvé			
IL18(2)	Aucun modèle approprié trouvé			
VWFT(1)	<i>age, BMI, Hchol_c, tabac_c, facteur 1, facteur 2</i>	0.011	40.72%	<i>BMI,</i>
VWFT(2)	<i>BMI, inh_all</i>	0.00147	29.06%	<i>BMI</i>
FGT(2)	Aucun modèle approprié trouvé			

Comme précédemment, on ne trouve pas de résultat concluant pour certains modèles, toujours pour les mêmes raisons, non respect des hypothèses sur les erreurs ou manque de données, voire encore modèle mauvais globalement.

Pour chaque modèle, les hypothèses sont bien vérifiées (on a procédé de la même manière que pour l'étude de la variable *ICAM(1)* dans la partie 2.1.1.1) et on a très peu de cas d'observations atypiques.

Nous allons donc vérifier que notre problème de multi-colinéarité a bien été résolu, en utilisant les VIFs. On a les valeurs de VIF suivantes, trouvées de la même manière que précédemment (matrice sweep courante en annexe_2):

Variables	Valeurs de VIF
<i>age</i>	1.39242
<i>sexe</i>	1.64731
<i>BMI</i>	2.10220
<i>EPS</i>	1.40810
<i>HTA</i>	10.1877
<i>TTT</i>	5.10190
<i>beta_c</i>	3.19378
<i>Inh_all</i>	3.07871
<i>Hchol_c</i>	2.43008
<i>Trigl_c</i>	1.45437
<i>tabac_c</i>	1.31984
<i>facteur 1</i>	3.13009
<i>facteur 2</i>	3.77190

On ne remarque qu'une seule valeur de VIF supérieure à 10 (et de très peu) pour la variable *HTA*, on en conclura que nous n'avons pas de problème de multi-colinéarité ici. Les résultats ont donc bien été améliorés grâce à l'ACP.

3 Différences entre le début et la fin de la nuit

Afin de voir s'il existe une différence significative entre le début et la fin de la nuit pour les variables endothéliales, nous allons utiliser une ANOVA à mesures répétées. Nous détaillerons la méthode utilisée uniquement pour la variable *ICAM* puis nous l'appliquerons à tous les autres facteurs endothéliaux.

3.1 Analyse détaillée pour le variable *ICAM*

3.1.1 Modèle

Nous voulons étudier l'effet d'une nuit de sommeil sur la protéine d'adhésion cellulaire *ICAM*. Nous disposons des mesures de la même variable prises avant et après la nuit sur les mêmes sujets, c'est pourquoi nous allons utiliser une ANOVA à mesures répétées, qui prend en compte la variabilité « inter-sujets » et la variabilité « intra-sujets ».

Le modèle s'écrit de la manière suivante :

$$Y_{is} = \mu + \alpha_i + \tau_s + \varepsilon_{is}$$

Où $i = 1, 2$, $s = 1, \dots, 42$, avec la contrainte :

$$\sum_{i=1}^2 \alpha_i = 0$$

Où Y_{is} est la valeur prise par la réponse Y dans les conditions (α_i, τ_s) .

Les conditions à vérifier sont les suivantes :

- Les effets aléatoires τ_s sont indépendants,
- Les effets aléatoires τ_s ont même variance,
- Les effets aléatoires τ_s sont de loi gaussienne.

- Les erreurs sont indépendantes,
- Les erreurs ont même variance,
- Les erreurs sont de loi gaussienne.

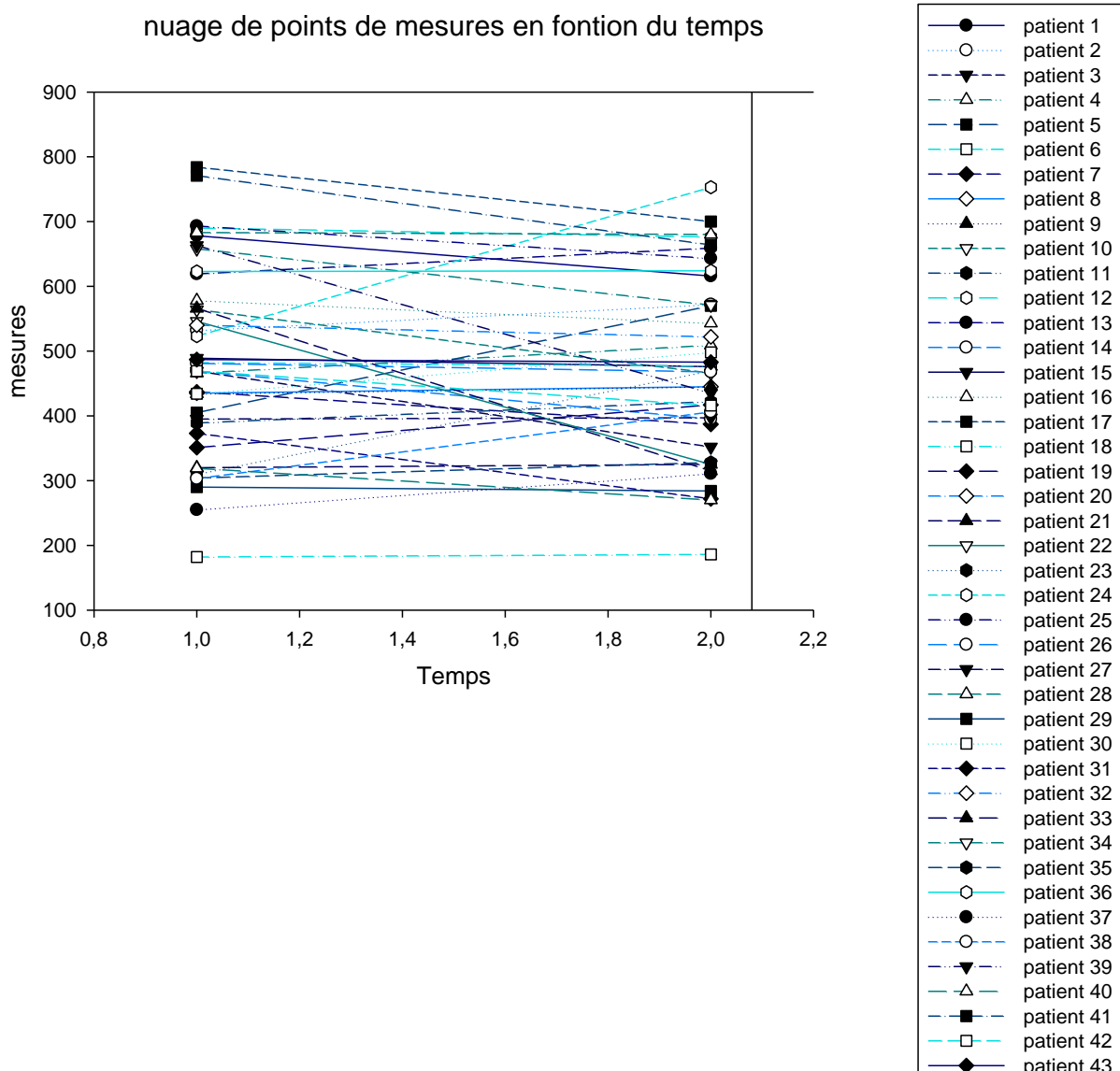
- Les effets aléatoires τ_s et les erreurs sont indépendants.

Nous allons donc commencer par vérifier les hypothèses.

3.1.2 Vérification des hypothèses

- Additivité du modèle

On trace le nuage de point de la réponse (mesure de l'adhésion cellulaire) en fonction du temps pour chaque patient, ce qui nous donne :



Etant donné que nous ne disposons que de deux temps différents (avant et après la nuit), il est difficile de conclure sur l'additivité du modèle.

- Indépendance du facteur τ

D'après le protocole expérimental, les sujets sont tous indépendants.

- Normalité du facteur τ

Il faudrait le vérifier en calculant et stockant les coefficients du modèle, nous n'avons malheureusement pas réussi à le faire avec le logiciel Statistica qui ne le permet pas.

- Homogénéité des variances du facteur τ

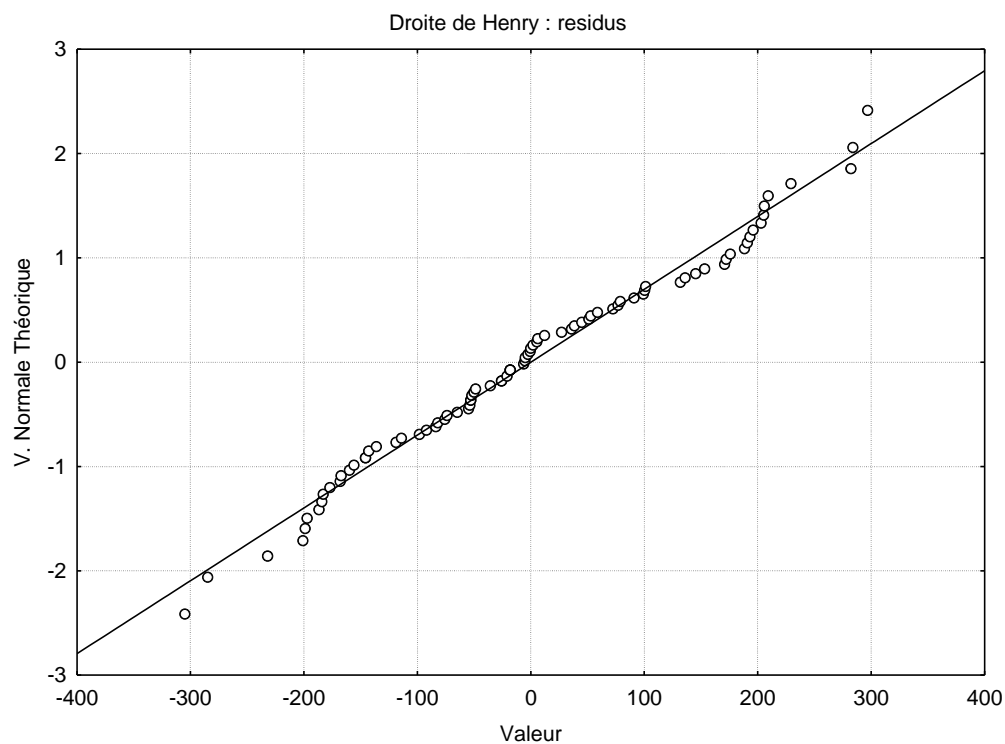
Nous ne pouvons pas évaluer cette hypothèse en l'absence de répétition pour chacun des patients.

- Indépendance des erreurs

Il n'existe pas de test.

- Normalité des erreurs

Afin de vérifier cette hypothèse, nous allons tracer la droite de Henry des résidus :



Les résidus sont bien alignés sur la droite, on peut en déduire que l'hypothèse de normalité des résidus est bien vérifiée.

- Homogénéité des variances des résidus

Etant donné que nous disposons de 42 observations et de seulement deux temps différents, il ne serait pas forcément intéressant de tracer le nuage de points des résidus en fonction du temps pour chaque patient. De plus, Statistica ne nous permet pas de la tester par rapport au facteur à effets fixes pour ce modèle.

3.1.3 Résultats

On veut tester l'hypothèse nulle

$$H_0 : \alpha_1 = \alpha_2 = 0$$

Contre l'hypothèse alternative

$$\exists i_0 \in \{1, 2\} : \alpha_{i_0} \neq 0$$

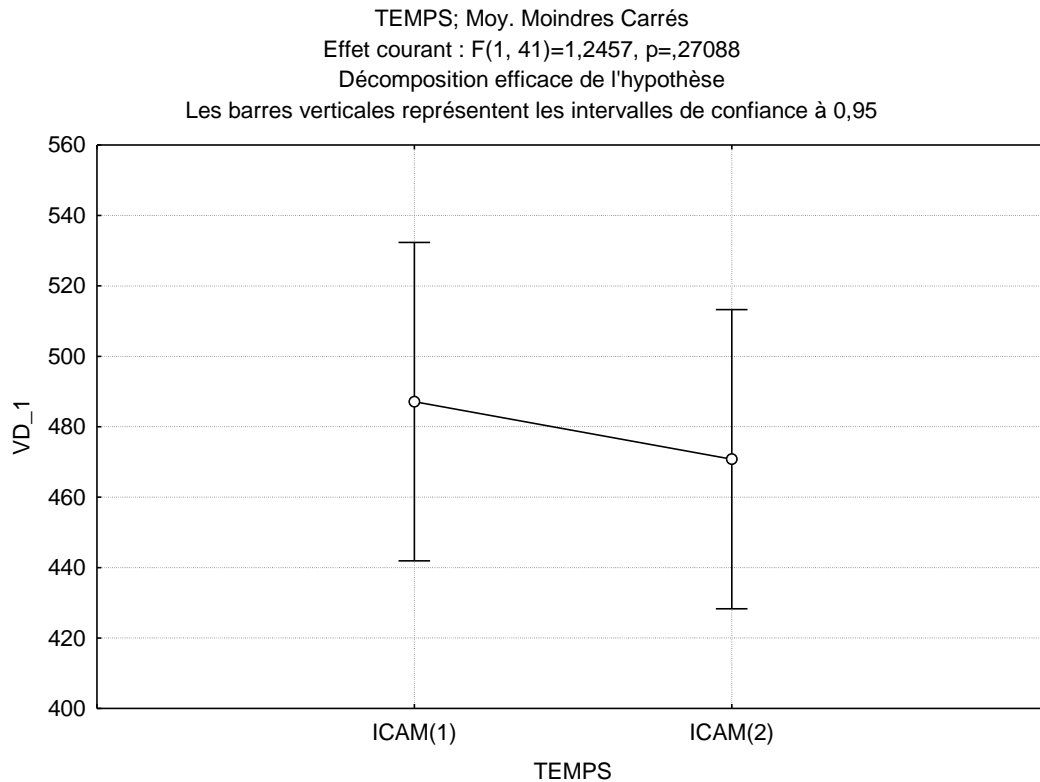
Sous l'hypothèse nulle et lorsque les conditions de validité du modèle sont respectées, $F_{\alpha, obs}$ est la réalisation d'une variable aléatoire qui suit une loi de Fisher à 1 et 41 degrés de liberté. On trouve les résultats suivants :

Analyse de la Variance Mesures Répétées (SAS + COR bio version 8_4_10_v5) Paramétrisation sigma-restreinte Décomposition efficace de l'hypothèse					
	SC	Degr. de - Liberté	MC	F	p
ord. origine	19269212	1	19269212	548,7837	0,000000
Erreur	1439616	41	35113		
TEMPS	5602	1	5602	1,2457	0,270880
Erreur	184396	41	4497		

On a la p-valeur associée égale à 0.270880 > 0.05 on ne peut donc pas rejeter l'hypothèse nulle. Le risque associé à cette décision est un risque de deuxième espèce, pour l'évaluer, il faudrait calculer la fonction puissance de ce test.

On en conclut donc qu'il n'y a pas d'effet significatif du temps (avant et après la nuit) sur l'adhésion cellulaire.

Afin d'illustrer cette conclusion, on trace le graphique correspondant :

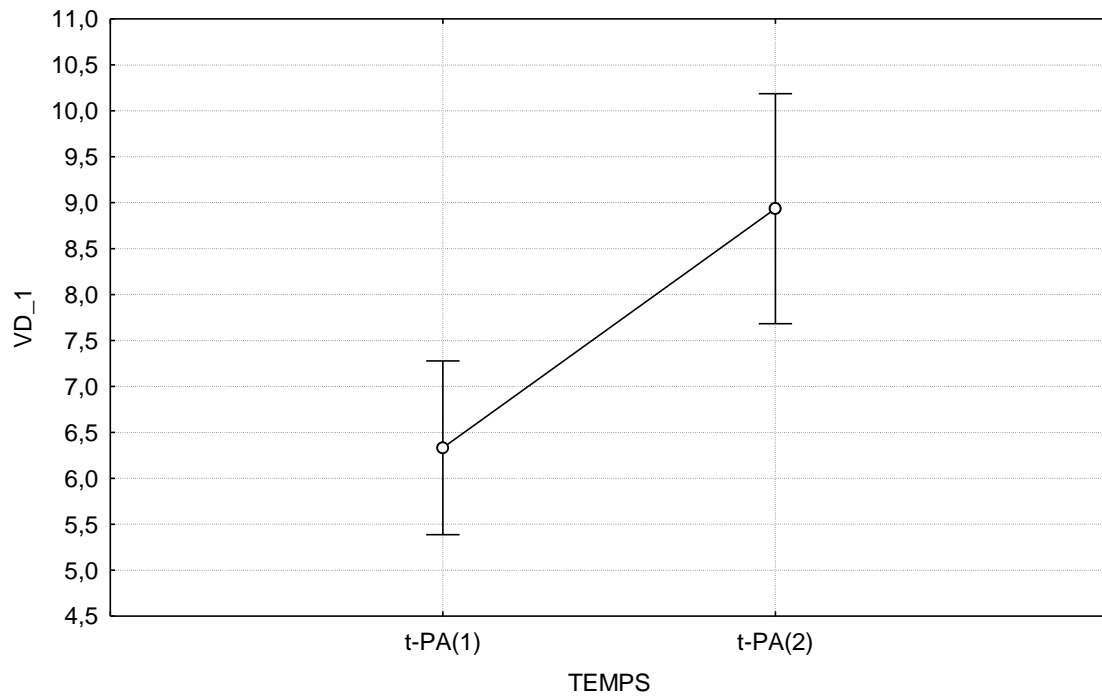


On effectue l'ANOVA à mesures répétées selon le même schéma pour tous les autres facteurs endothéliaux. On ne présentera qu'une synthèse générale des résultats trouvés.

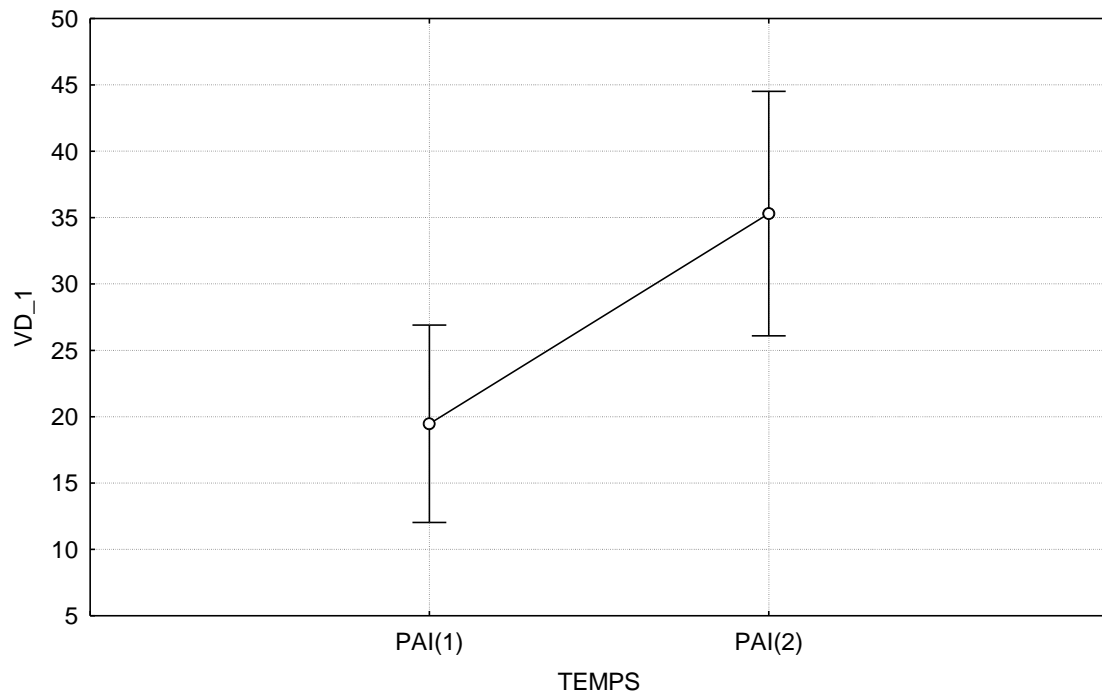
3.2 Résultats des analyses pour les autres facteurs endothéliaux

Après avoir effectué l'analyse suivant le même modèle que précédemment, on trouve un effet significatif du temps (avant et après la nuit) sur les facteurs endothéliaux suivants : le plasminogène tissulaire (*t-PA*), et l'inhibiteur de l'activateur du plasminogène (*PAI*). Pour ces deux facteurs, on trouve les graphiques suivants :

TEMPS; Moy. Moindres Carrés
Effet courant : $F(1, 42)=44,209$, $p=,00000$
Décomposition efficace de l'hypothèse
Les barres verticales représentent les intervalles de confiance à 0,95



TEMPS; Moy. Moindres Carrés
Effet courant : $F(1, 42)=13,740$, $p=,00061$
Décomposition efficace de l'hypothèse
Les barres verticales représentent les intervalles de confiance à 0,95



On peut également remarquer que les deux variables évoluent dans le même sens, ce qui semble logique. La plasminogène est une enzyme initiant la fibrinolyse (processus physiologique complexe de dissolution des caillots sanguins) et l'activateur tissulaire du plasminogène est inhibé en fin de fibrinolyse par les inhibiteurs de l'activateur du plasminogène (PAI). Il semble donc normal que quand la première augmente, la seconde augmente de la même manière, afin d'éviter que la première augmente encore plus et de retrouver un équilibre.

Maintenant que nous avons regardé les effets de la nuit de sommeil sur les facteurs endothéliaux, nous allons nous intéresser aux effets des différents traitements sur les variables du sommeil et les facteurs endothéliaux.

4 Effets des traitements

Afin de se rendre compte s'il existe une différence significative due aux traitements administrés aux patients souffrant de troubles du sommeil, nous allons effectuer, comme précédemment, une ANOVA à mesures répétées. Nous le ferons indépendamment pour le traitement par PPC et pour le traitement par O2.

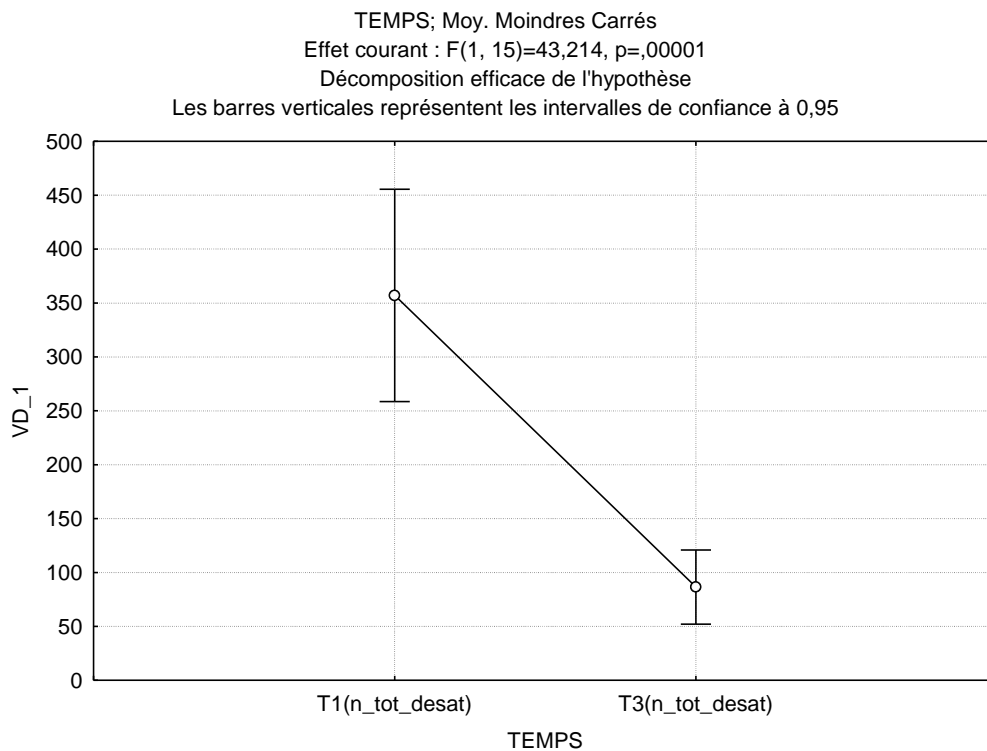
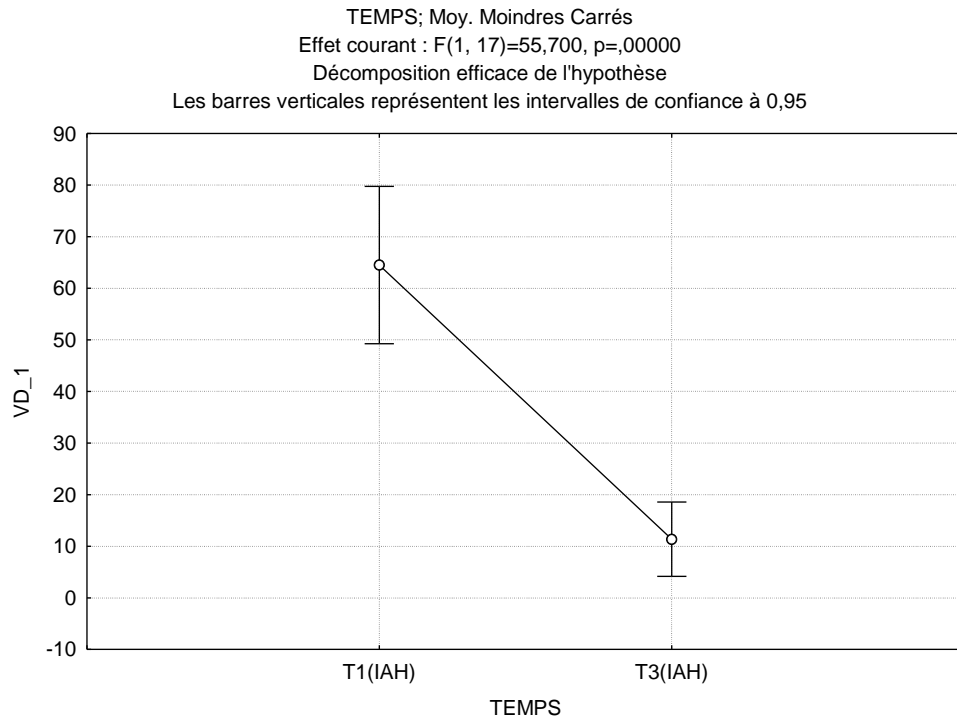
4.1 Traitement par PPC

Nous disposons de données pour des patients après une nuit et après un mois de traitement PPC. Nous avons ainsi 19 observations pour une nuit de traitement et, à cause de la difficulté du suivi à long terme des patients, de seulement 11 observations pour un mois de traitement. Afin de tenir compte de toutes les observations disponibles pour le traitement d'une nuit, nous effectuerons tout d'abord une ANOVA à mesures répétées pour voir les effets de ce traitement, puis nous referons l'ANOVA à mesures répétées avec une nuit et un mois de traitement.

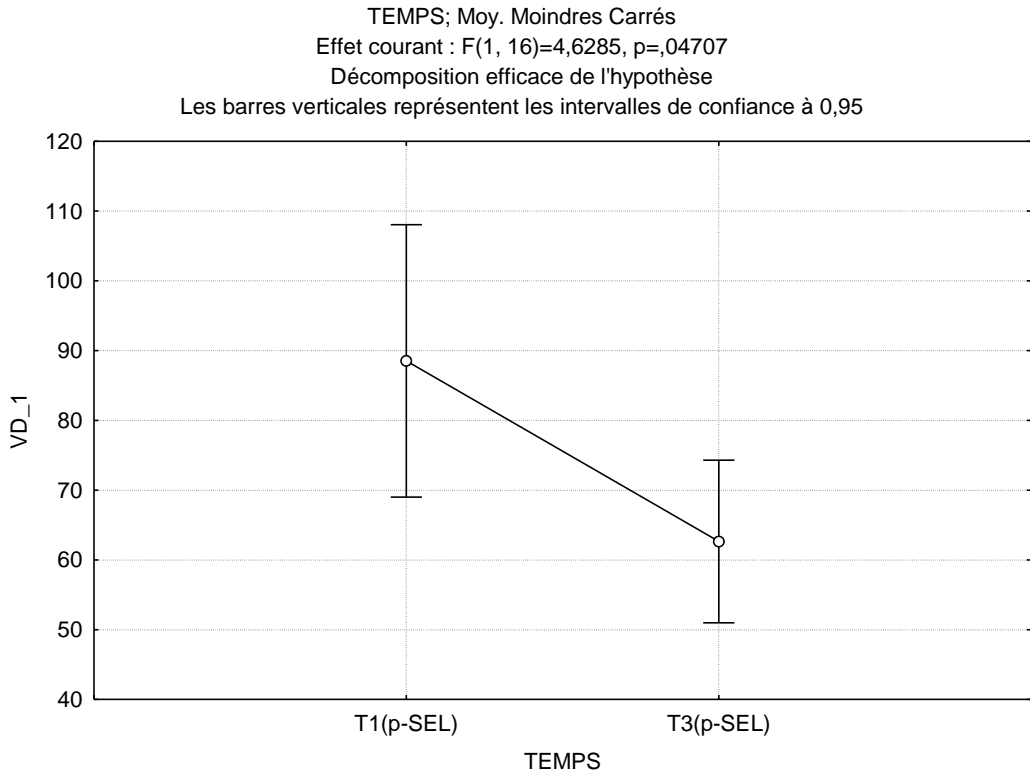
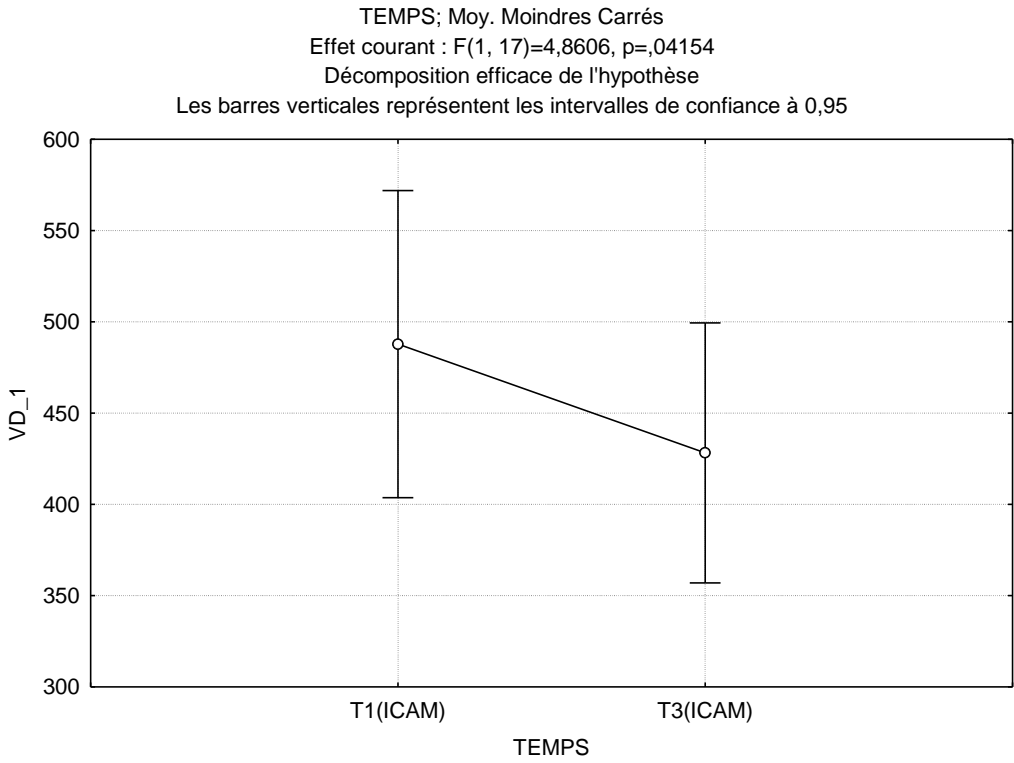
4.1.1 Effet d'une nuit de traitement par PPC

Nous ne donnerons pas ici le détail sur le modèle et sur la vérification des hypothèses car cela a été effectué suivant le même principe que dans la partie 3.1.

Les variables du sommeil et conséquences des troubles du sommeil pour lesquelles nous disposons de données suffisantes sont les variables *IAH*, *TTS*, *Eff_som%*, *SLP%*, *SP%*, *V_min*, *sat_moy%* et *n_tot_desat*. Nous avons trouvé un effet significatif (au seuil de 5%) de la nuit de traitement par PPC sont l'indice apnée hypopnée (*IAH*), et le nombre total de désaturations (*n_tot_desat*). Pour ces variables, on trouve les graphiques associés suivants :



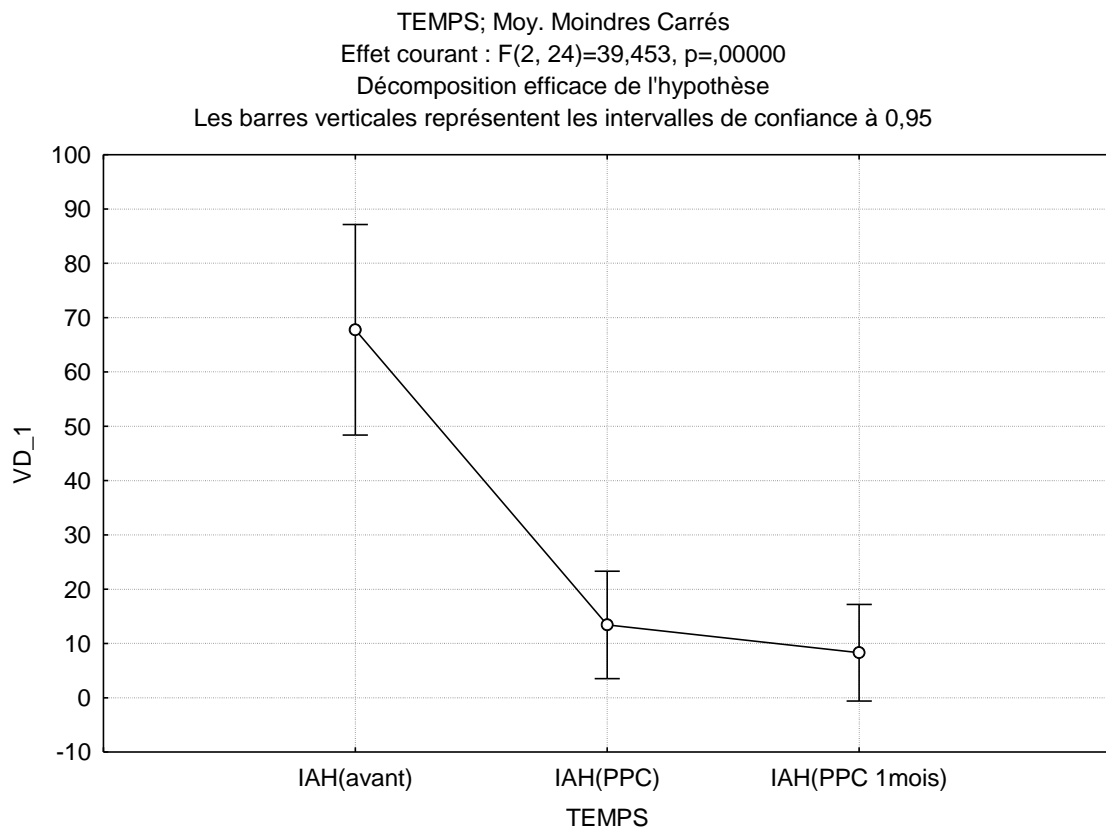
On trouve également un effet significatif (au seuil de 5%) sur les facteurs endothéliaux suivants : l'adhésion cellulaire mesurée le soir (*ICAM(1)*) et la sélectine plaquettaire mesurée le soir (*p-SEL(1)*). Voici les graphiques correspondants :



4.1.2 Effet d'un mois de traitement PPC

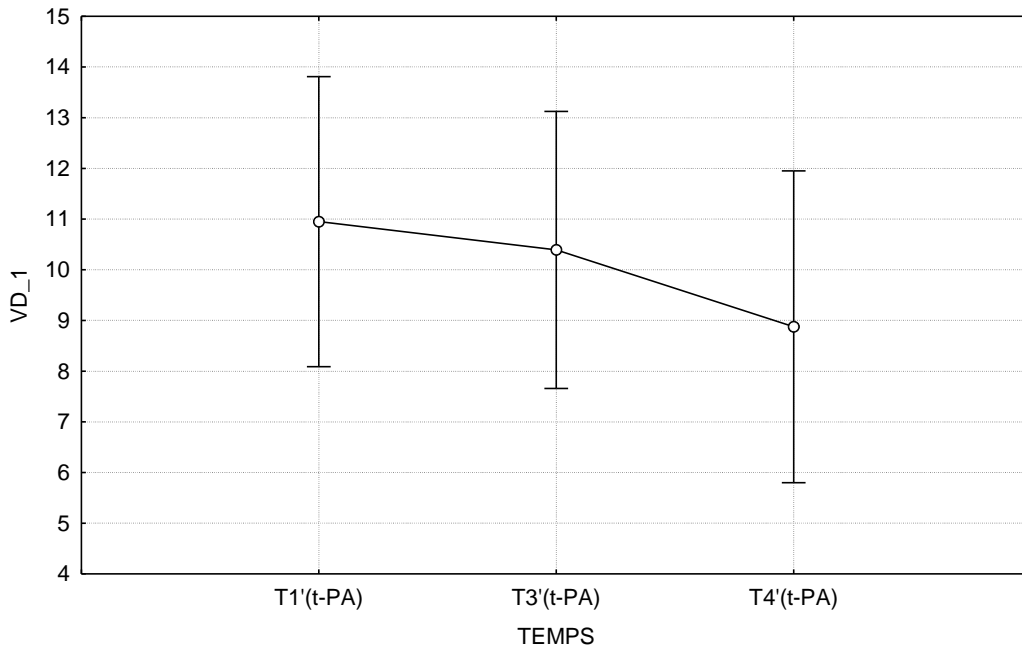
Nous allons faire une ANOVA à mesures répétées pour chaque variable dont nous disposons de données suffisantes, c'est-à-dire *IAH*, *sat_moy%* et toutes les variables endothéliales. Nous ne détaillerons pas le modèle et la vérification des hypothèses, mais elle est faite suivant le même principe que dans la partie 3.1.

On trouve un effet significatif (au seuil 5%) du traitement sur l'indice d'apnée hypopnée (*IAH*), qui nous donne le graphique suivant :

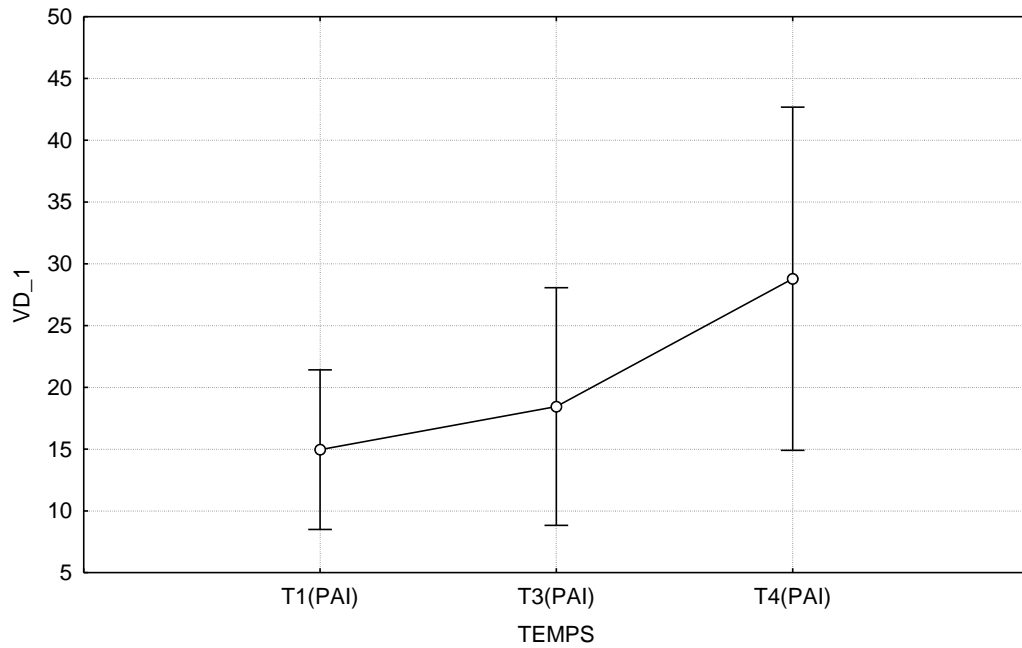


Pour les facteurs endothéliaux, on trouve un effet significatif (au seuil 5%) du temps pour le tissu plasminogène mesuré le matin (*t-PA(2)*) et pour l'inhibiteur de l'activateur de plasminogène mesuré le soir (*PAI(1)*). On a les graphiques correspondant suivants :

TEMPS; Moy. Moindres Carrés
 Effet courant : $F(2, 22)=5,7746$, $p=,00964$
 Décomposition efficace de l'hypothèse
 Les barres verticales représentent les intervalles de confiance à 0,95



TEMPS; Moy. Moindres Carrés
 Effet courant : $F(2, 20)=4,9316$, $p=,01815$
 Décomposition efficace de l'hypothèse
 Les barres verticales représentent les intervalles de confiance à 0,95



Remarques :

- On peut d'abord remarquer que nous avons toujours un effet significatif du traitement au bout d'un mois pour l'indice d'apnée hypopnée (*IAH*)
- On peut également remarquer que nous ne trouvons pas d'effet significatif d'un mois de traitement pour les variables *ICAM(1)* et *p-SEL(1)*, alors que l'on avait trouvé un effet significatif d'une nuit de traitement pour celles-ci. Cela peut venir du fait que nous avons moins de patients suivis sur un mois que sur une nuit de traitement.
- On peut enfin faire remarquer que nous trouvons un effet significatif d'un mois de traitement sur le tissu plasminogène mesuré le soir (*t-PA(1)*), ainsi que sur l'inhibiteur de l'activateur plasminogène mesuré le matin (*PAI(2)*).

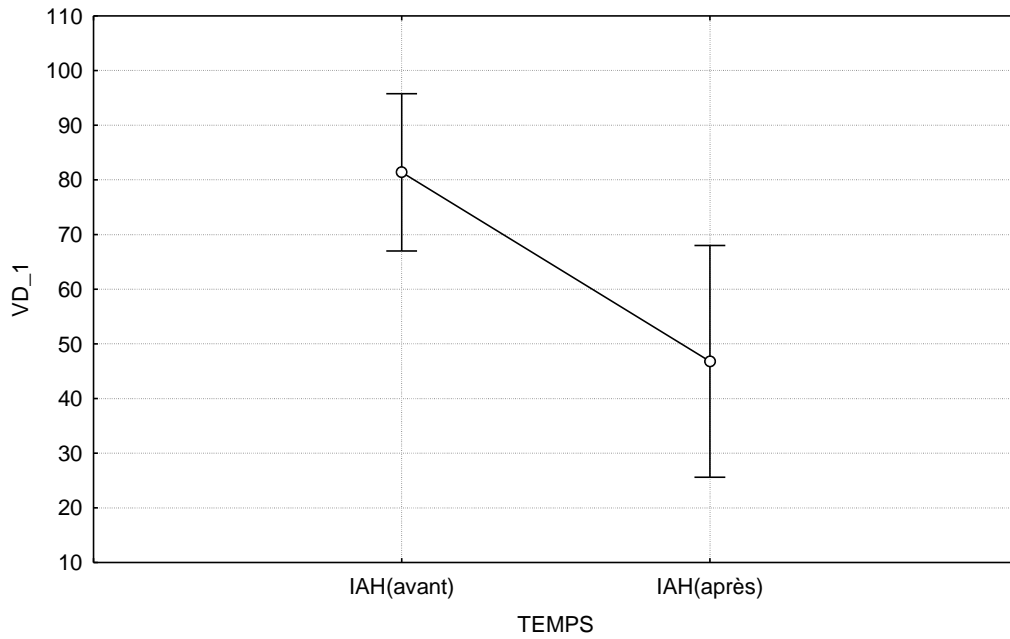
4.2 Traitement par O2

Nous disposons de données pour 11 patients suivis après une nuit de traitement par O2. Afin d'étudier les effets de ce traitement, nous allons à nouveau effectuer une ANOVA à mesures répétées. Comme précédemment, nous ne donnerons pas ici le détail sur le modèle et sur la vérification des hypothèses car cela a été effectué suivant le même principe que dans la partie 3.1.

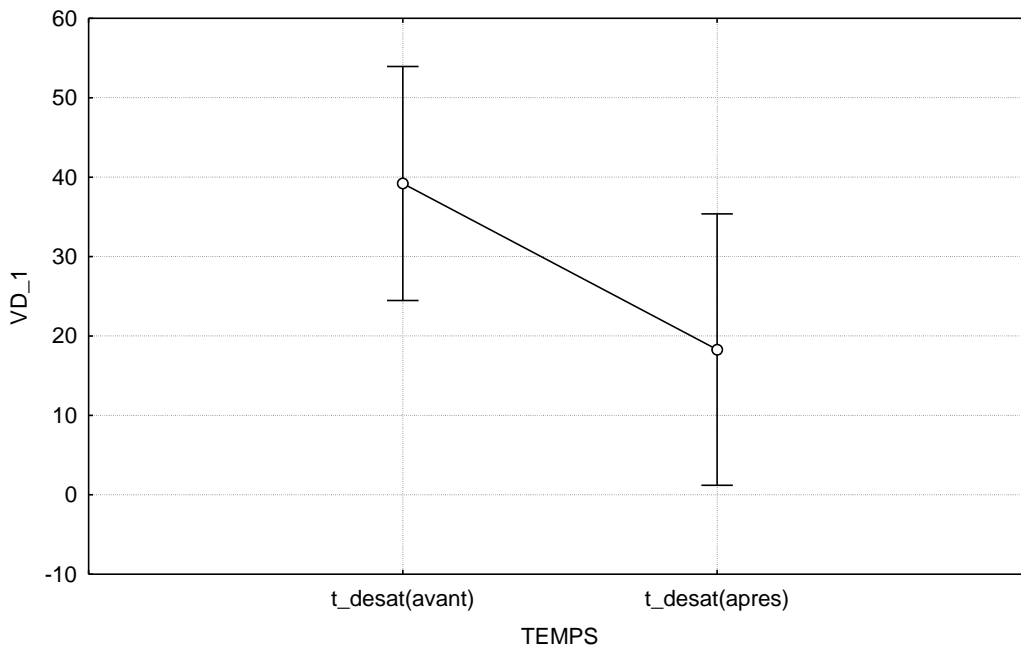
On a réalisé l'ANOVA pour les variables du sommeil et conséquences des troubles du sommeil *IAH*, *TTS*, *Eff_som%*, *SLP%*, *SP%*, *V_min*, *t_desat%*, *sat_moy%* et *n_tot_desat*, ainsi que pour tous les facteurs endothéliaux.

On a trouvé un effet significatif d'une nuit de traitement par O2 pour indice d'apnée hypopnée (*IAH*), et les variables de saturation en oxygène *t_desat*, *sat_moy%* et *n_tot_desat*. On a les graphiques correspondant :

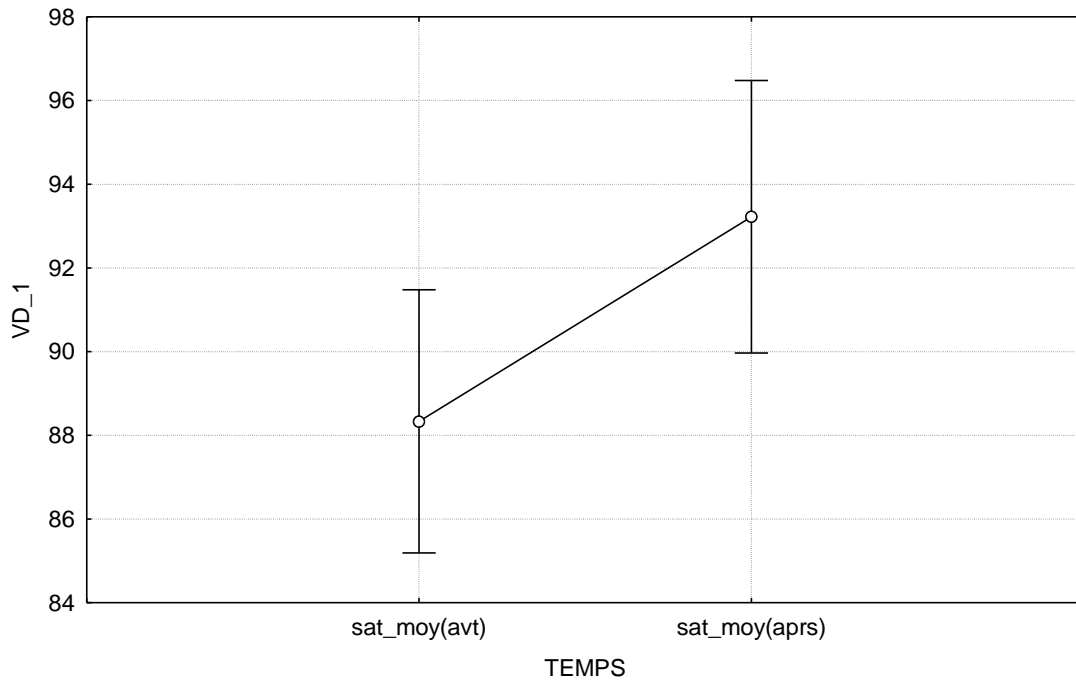
TEMPS; Moy. Moindres Carrés
Effet courant : $F(1, 9)=10,164$, $p=,01104$
Décomposition efficace de l'hypothèse
Les barres verticales représentent les intervalles de confiance à 0,95



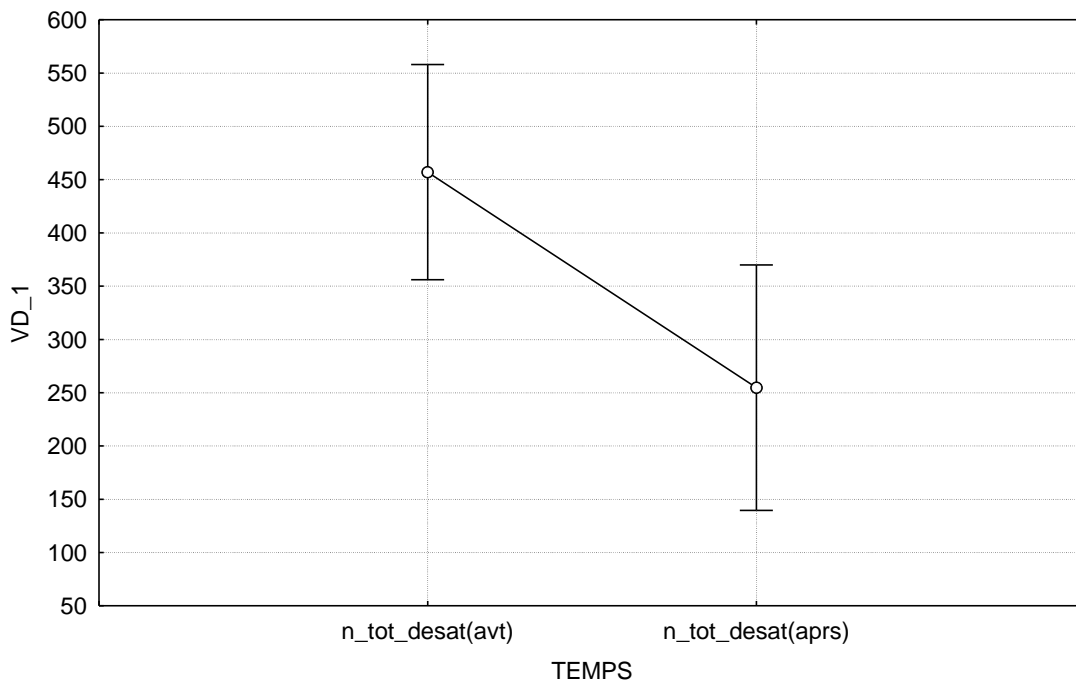
TEMPS; Moy. Moindres Carrés
Effet courant : $F(1, 8)=33,030$, $p=,00043$
Décomposition efficace de l'hypothèse
Les barres verticales représentent les intervalles de confiance à 0,95



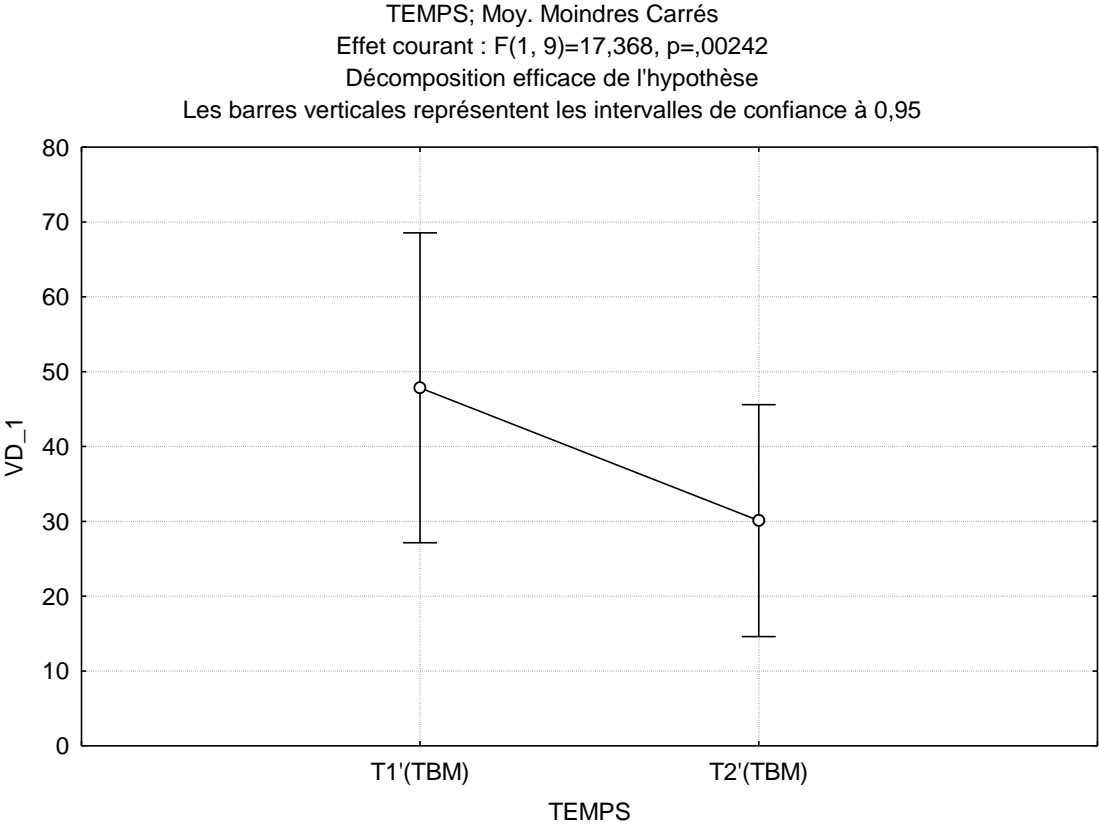
TEMPS; Moy. Moindres Carrés
 Effet courant : $F(1, 8)=20,761$, $p=,00186$
 Décomposition efficace de l'hypothèse
 Les barres verticales représentent les intervalles de confiance à 0,95



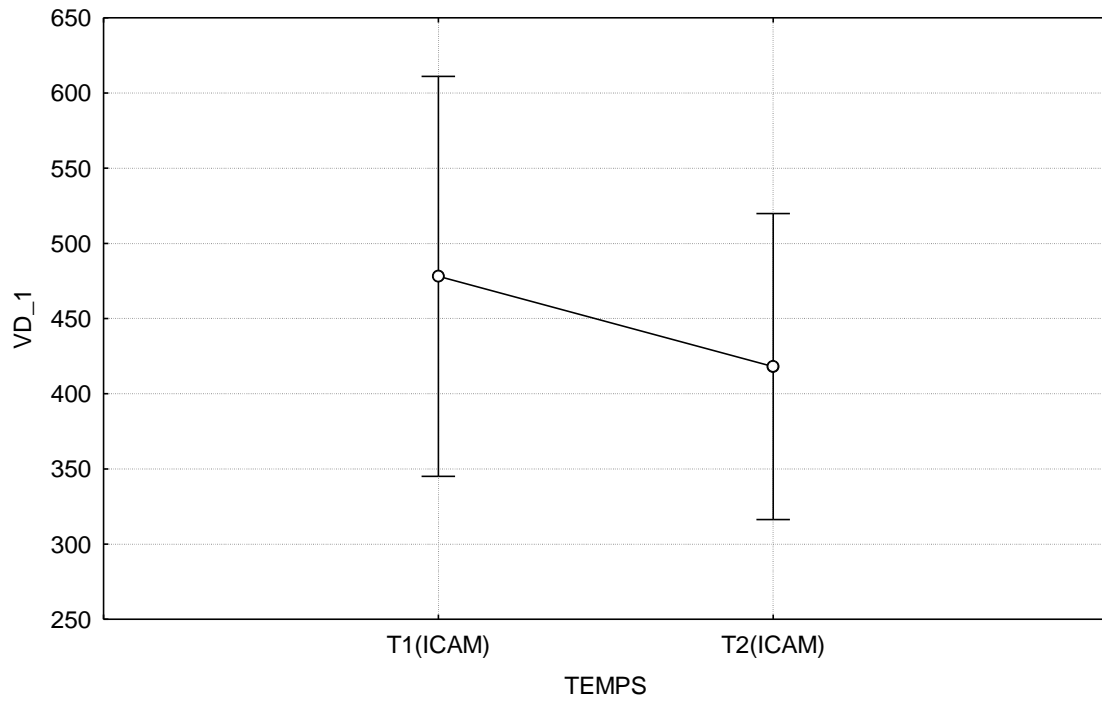
TEMPS; Moy. Moindres Carrés
 Effet courant : $F(1, 8)=26,483$, $p=,00088$
 Décomposition efficace de l'hypothèse
 Les barres verticales représentent les intervalles de confiance à 0,95



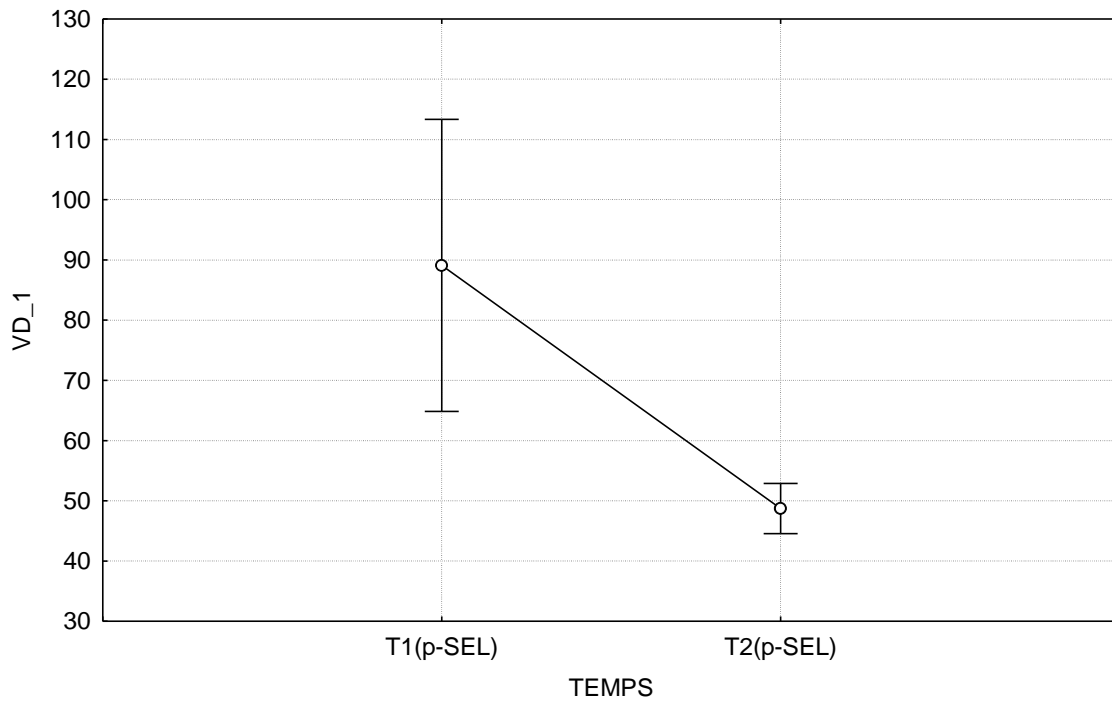
Pour les facteurs endothéliaux, on a trouvé un effet significatif d'une nuit de traitement par O2 pour la thrombomoduline mesurée le matin (*TBM(2)*), l'adhésion cellulaire mesurée le soir (*ICAM(1)*) et la sélectine plaquettaire mesurée le soir (*p-SEL(1)*). On a les graphiques correspondant :



TEMPS; Moy. Moindres Carrés
Effet courant : $F(1, 10)=5,6815$, $p=,03838$
Décomposition efficace de l'hypothèse
Les barres verticales représentent les intervalles de confiance à 0,95



TEMPS; Moy. Moindres Carrés
Effet courant : $F(1, 10)=14,596$, $p=,00337$
Décomposition efficace de l'hypothèse
Les barres verticales représentent les intervalles de confiance à 0,95



Remarques :

- On peut tout d'abord remarquer que nous trouvons un effet significatif d'une nuit de traitement par O2 et d'une nuit de traitement par PPC sur l'indice d'apnée hypopnée et sur les variables de saturation en oxygène.
- On remarque également que l'on retrouve un effet significatif pour les traitements d'une nuit PPC et O2 sur l'adhésion cellulaire et la sélectine plaquettaire, tandis que le traitement sur un mois de PPC a un effet significatif sur le tissu plasminogène et l'inhibiteur de l'activateur plasminogène.
- On peut faire une dernière remarque sur le fait que nous avons trouvé un effet sur la thrombomoduline que nous n'avions pas avec le traitement par PPC.

En conclusion, il semble que le traitement à court terme, que ce soit par PPC ou oxygénothérapie a une répercussion sur les facteurs d'adhésion cellulaire et de sélectine plaquettaire, alors que le traitement à long terme par PPC a plutôt une répercussion sur le tissu plasminogène.

5 Conclusion

L'ensemble de cette étude a permis de mettre en évidence un lien entre les troubles respiratoires du sommeil et les facteurs endothéliaux. Nous avons ainsi montré que le fait de traiter ces problèmes d'apnées hypopnées, que ce soit par pression positive continue ou par oxygénothérapie, a une répercussion directe sur certains de ces facteurs. Plus précisément, un traitement à court terme par PPC ou oxygénothérapie (même si dans ce cas, le sommeil reste fragmenté) va agir sur l'adhésion cellulaire et les plaquettes, tandis qu'un traitement à long terme (uniquement par PPC ici) aura plutôt une influence sur le tissu plasminogène, impliqué dans la coagulation sanguine. De plus, la dysfonction endothéliale pouvant être un facteur de risques de complications cardiovasculaires, le fait de mieux comprendre le lien entre les troubles respiratoires du sommeil et les facteurs endothéliaux permettra une meilleure prise en charge de ces risques chez les patients sujets aux apnées du sommeil.

Durant ce stage, j'ai pu appliquer les méthodes statistiques vues en cours à des données réelles, et ainsi comprendre leur utilité tout en approfondissant mes connaissances. J'ai également pu me familiariser avec le logiciel Statistica, souvent utilisé par les biologistes. De plus, ce stage ayant été réalisé en interaction avec des cliniciens et des chercheurs, cela m'a permis de comprendre l'importance des statistiques dans le domaine de la biologie et de la recherche. Cette expérience professionnelle a donc été très enrichissante et m'a donné une nouvelle vue du métier de statisticien, ce qui me permettra de mieux m'orienter professionnellement.

6 Annexe_1

Etat courant de Matrice Sweep : VD-t(PA(1)) (SAS + COR bio version 8.4.10 w4)

	age	sexe	BMI	EPS	HTA	TTC	beta_c	inh_All	Hehol_c	Tripl_c	Tabac_c	IAH	n_tot_ewt	ILD_h	Diag_t4	TTS	Eff_som%	R %	SLP%	SP%	V_min	thesat%	sat_moy %	n_tot_desat	n_desat_h	traimt	
age	7.89	0.22	-3.81	1.63	6.63	2.13	-0.00	-7.20	8.46	-0.37	-2.60	3.70	1.78	1.13	5.73	12.94	36.67	-3.81	1.41	6.78	34.54	1.49	-5.92	1.29	-14.12	0.86	
sexe	0.22	4.55	2.28	-0.22	2.88	1.73	-2.23	-2.81	2.17	-0.28	-0.28	5.35	4.79	-1.37	-6.80	-2.97	8.69	-1.03	0.18	1.73	9.85	1.30	4.63	-1.33	-2.32	-1.45	
BMI	-3.81	2.28	4.55	2.28	2.60	0.20	0.93	-3.02	3.79	-0.39	-0.73	-0.34	-1.17	0.03	7.82	-4.11	13.87	-1.09	0.70	-3.23	13.04	-2.22	-9.01	-0.33	-6.65	0.59	
EPS	1.63	-0.22	2.28	4.55	4.49	-0.27	2.19	3.01	-2.90	0.38	0.62	-11.09	6.57	-2.55	-6.47	0.21	-8.82	2.97	-0.42	-0.84	-9.12	5.71	8.72	6.32	4.05	-0.75	
HTA	6.63	2.28	2.60	4.49	23.01	10.47	7.44	15.27	-13.93	0.77	2.82	-28.80	5.49	-4.06	-6.47	12.11	-44.29	7.40	-2.51	-8.56	-44.69	2.84	7.20	-0.16	23.88	-0.15	
TTC	2.13	1.73	0.20	-0.27	10.47	7.65	3.42	3.42	-5.29	3.31	-3.31	5.78	-2.50	3.25	4.66	3.36	-13.90	1.66	-0.96	-3.33	-14.23	-1.66	-1.10	0.28	5.78	0.37	
beta_c	-0.00	-2.23	0.93	2.19	7.44	3.42	-5.29	-3.31	3.31	2.50	0.04	11.29	-1.13	0.47	-2.42	0.36	7.29	-1.81	-0.06	0.64	8.39	-3.00	-0.95	-4.68	-2.76	-0.19	
inh_All	-7.20	-2.81	-3.02	3.01	15.27	5.78	-3.31	-17.57	15.93	0.06	-4.33	33.59	-9.44	6.78	-1.94	-16.71	54.58	-8.68	3.75	11.40	54.99	1.39	-4.71	4.22	-27.31	-1.25	
Hehol_c	8.46	2.17	3.79	-2.90	13.93	-5.59	2.50	15.93	-17.01	0.05	4.70	-31.63	9.25	-7.84	0.00	20.60	-57.94	8.40	-3.50	-11.20	-57.66	-2.50	6.08	-7.25	30.58	0.79	
Tripl_c	-0.37	-0.28	-0.39	0.38	0.77	0.61	-0.49	0.06	0.05	-2.02	0.33	-7.62	7.08	-4.19	0.48	1.15	1.12	1.21	-0.06	-0.20	0.34	-1.46	-1.11	-6.00	4.97	0.64	
Tabac_c	-2.60	-0.28	-0.73	0.62	2.82	1.38	0.04	-4.33	4.70	0.33	4.73	6.99	-4.51	3.50	0.85	-8.42	8.38	-2.69	1.66	4.57	9.14	3.28	1.20	10.27	-12.80	-1.40	
IAH	3.70	5.35	-0.34	11.09	28.80	-7.68	11.29	33.59	-31.63	-7.62	6.99	-442.19	325.42	-287.27	125.64	24.40	41.42	-121.63	29.05	-6.51	-7.07	-125.97	-7.44	1.20	-78.27	143.66	-0.90
n_tot_ewt	1.78	4.79	-1.17	6.57	5.49	-1.26	-1.13	-8.44	9.25	7.08	-4.51	325.42	-287.27	125.64	1.36	-18.82	31.10	-14.48	1.94	-5.44	33.06	7.89	-2.80	84.08	-99.64	-0.90	
ILD_h	1.13	-1.37	0.03	-2.55	-4.06	-0.74	0.47	6.78	-7.84	4.19	3.50	-146.04	125.64	-69.22	-0.47	14.41	-26.67	9.19	-2.01	0.03	-28.33	-7.14	0.21	-42.12	53.20	0.18	
Diag_t4	5.73	-6.80	7.82	-6.47	0.41	3.25	-2.42	-1.94	0.00	0.48	0.86	24.40	1.36	-0.47	-34.30	1.14	-17.35	2.50	-0.09	-0.10	-13.05	9.45	24.21	6.98	3.34	-2.30	
TTS	12.94	-2.97	4.11	0.21	12.11	4.66	0.36	-18.71	20.60	1.15	-8.42	41.42	-18.82	14.41	1.14	47.70	118.99	-11.12	4.35	16.65	109.19	15.97	1.75	34.62	-60.51	-1.87	
Eff_som%	36.67	8.69	13.87	-8.62	44.29	13.90	7.29	54.58	-57.94	1.12	8.38	-121.63	31.10	-26.67	-17.35	118.99	-445.23	34.68	-10.14	-39.48	-403.63	-14.61	25.75	1.27	120.27	-5.49	
R %	-3.81	-1.03	-1.09	2.97	7.40	1.86	-1.81	-8.68	8.40	1.21	-2.69	29.05	-14.48	9.19	2.50	-11.12	34.68	-8.08	2.09	5.69	34.86	-1.02	-5.73	5.37	-21.63	-1.04	
SLP%	1.41	0.18	0.70	-0.42	-2.51	-0.96	-0.06	3.75	-3.50	-0.06	1.66	-6.51	1.94	-2.01	-0.09	4.35	-10.14	2.09	-2.38	-2.81	-10.82	-0.87	1.03	-1.74	6.23	0.88	
SP%	6.78	1.73	3.23	-0.84	-8.56	-3.33	0.64	11.40	-11.20	-0.20	4.57	-7.07	-5.44	0.03	-0.10	16.65	-39.48	5.69	-2.81	-11.50	-39.20	-3.29	2.24	-3.47	17.16	1.71	
V_min	34.54	9.85	13.04	-9.12	44.69	14.23	8.39	54.99	-47.69	0.34	9.14	-125.97	33.06	-13.05	-10.82	-39.20	-403.63	34.86	-10.82	-39.20	-12.35	-23.57	-22.14	-29.71	19.46	1.87	
thesat%	1.49	1.30	-2.22	5.71	2.84	-1.66	-3.00	1.39	-2.50	-1.46	3.28	-7.44	7.89	-7.14	9.45	15.97	-14.61	-1.02	-0.87	-3.29	-12.35	-23.57	-22.14	-29.71	19.46	1.87	
sat_moy %	-5.92	4.63	-9.01	8.72	7.20	-1.10	-0.95	4.71	6.08	-4.22	4.22	4.71	-2.80	25.14	24.21	1.75	25.75	-5.73	1.03	2.24	25.14	-22.14	-35.79	-20.43	-150.22	130.59	8.04
n_tot_desat	1.29	-1.33	-0.33	6.32	-0.16	0.28	-4.68	4.22	-7.25	-6.00	10.27	-78.27	84.08	-42.12	6.98	34.62	1.27	5.37	-1.74	-3.47	-5.12	-29.71	-4.72	-130.59	-175.22	-6.38	
n_desat_h	14.12	-2.32	-6.65	4.05	23.88	5.78	-2.76	-27.31	30.58	4.97	-12.80	143.66	-99.64	52.20	3.34	-50.51	120.27	-21.63	6.23	17.16	122.25	-2.71	8.04	-6.38	-3.56		
traimt	0.86	-1.45	0.59	-0.75	-0.15	0.37	-0.19	-1.25	0.79	0.64	-1.40	-0.90	1.95	0.18	-2.30	-1.87	-5.49	-1.04	0.88	1.71	-2.71	1.87	2.72	8.04	-6.38	-3.56	

Matrice sweep courante (partie 2.1.3)

7 Annexe_2

Etat courant de Matrice Sweep ; VD: FGT(2) (SAS + COR bio version 8_4_10 v5)													
	age	sexe	BMI	EPS	HTA	TTT	beta_c	inh_All	Hchol_c	Trigl_c	Tabac_c	facteur 1	facteur 2
age	-1,39282	0,03199	-0,12895	-0,01827	0,8619	0,29074	-0,10123	-0,09035	0,34144	-0,34206	0,17584	-0,00336	-0,08717
sexe	0,03199	-1,64731	0,86461	0,20411	0,3378	0,34858	-0,56260	0,11075	-0,04523	-0,33959	-0,05577	-0,10811	-0,81540
BMI	-0,12895	0,86461	-2,10220	0,42997	-0,0812	-0,02340	0,45766	-0,30287	0,33410	0,11558	0,36807	0,16302	1,06269
EPS	-0,01827	0,20411	0,42997	-1,40810	-0,8392	-0,31587	0,66320	0,23802	-0,27392	0,13680	-0,05239	-0,29421	0,20123
HTA	0,86187	0,33777	-0,08124	-0,83920	-10,1877	-6,05732	4,09861	3,19629	-2,14309	0,95096	0,28905	0,05282	0,18030
TTT	0,29074	0,34858	-0,02340	-0,31587	-6,0573	-5,10190	2,24636	1,55148	-1,53656	0,74199	0,42627	0,08668	-0,05542
beta_c	-0,10123	-0,56260	0,45766	0,66320	4,0986	2,24636	-3,19378	-0,78351	0,51680	-0,38879	-0,39063	0,52993	-1,05342
inh_All	-0,09035	0,11075	-0,30287	0,23802	3,1963	1,55148	-0,78351	-3,07871	1,71866	-0,55881	-0,02710	-0,11517	0,75557
Hchol_c	0,34144	-0,04523	0,33410	-0,27392	-2,1431	-1,53656	0,51680	1,71866	-2,43008	0,86137	0,09634	0,51950	-0,84472
Trigl_c	-0,34206	-0,33959	0,11558	0,13680	0,9510	0,74199	-0,38879	-0,55881	0,86137	-1,45437	0,03379	-0,21984	-0,02961
Tabac_c	0,17584	-0,05577	0,36807	-0,05239	0,2891	0,42627	-0,39063	-0,02710	0,09634	0,03379	-1,31984	0,41857	-0,18080
facteur 1	-0,00336	-0,10811	0,16302	-0,29421	0,0528	0,08668	0,52993	-0,11517	0,51950	-0,21984	0,41857	-3,13009	2,30695
facteur 2	-0,08717	-0,81540	1,06269	0,20123	0,1803	-0,05542	-1,05342	0,75557	-0,84472	-0,02961	-0,18080	2,30695	-3,77190

Matrice sweep courante (partie 2.3)

8 Sources

Mémoire pour l'obtention du D.E.S de Neurologie Interrégion du Nord-Est, E.RUPPERT, 2009

<http://inci.u-strasbg.fr/fr/index.html>

<http://www.caraibesleepcenter.com/TROS>

http://www-irma.u-strasbg.fr/~fbertran/enseignement/DataMining_2009/ACP.pdf

http://fr.wikipedia.org/wiki/Droite_de_Henry