

Analyse exploratoire par technique de « Data Mining » des données recueillies à l'occasion de dix ans de pratique de médecine du sommeil.

Didier Cugy ^{1, 2}

(1) Unité d'Hypnologie - Pavillon de la Mutualité Pessac, (2) Clinique du Sommeil – SEFSN - CHU Bordeaux

Introduction :

La mise en place d'un dossier médical informatisé architecturé autour d'une base de données intégrant toutes les informations issues de la pratique de médecine du sommeil (Données biologiques, morphométriques, polysomnographies, questionnaires, examen clinique) permet d'agréger ces informations et de les exploiter par la technique du "data-mining"(1,2) .

Matériel et Méthodes :

Application de technique de DataMining a une base de données constituée en dix ans d'activité dédiée à la médecine du sommeil. La base de données intègre l'ensemble des données médico-administratives des patients, les examens cliniques, les réponses aux auto-questionnaires, les courriers, les enregistrements polysomnographiques (données de synthèses et données brutes), résultats des examens biologiques dont phénotypages. Les données des autoquestionnaires, des examens cliniques, biologiques et polysomnographiques réalisés à une même période sont associées pour l'analyse.

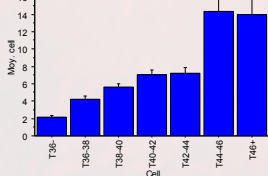
Résultats :

La base de données constituée a une taille de 489 GigaOctets et inclus 2323 dossiers avec données d'enregistrements polysomnographiques. Il est retrouvé de nombreuses associations connues comme Tour de Cou et index d'apnées-hypopnées ($p < 0.001$) (fig1,2), index de désaturation et score d'Epworth ($p < 0.03$)(3). De façon plus surprenante il est identifié une relation significative entre le statut en vitamine D et la latence d'endormissement ($p=0.04$)(fig3) ainsi que la concentration en folates (fig4), la période totale de sommeil et la durée du stade N1 ($p=0.04$), le temps total de sommeil et le score d'Epworth.

Graphique des interactions pour ID

Effet : CAT_TDC

Barres d'erreur: ± 1 Erreur(s) standard



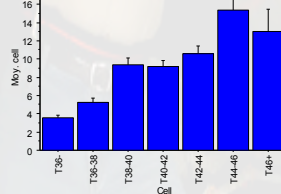
CAT_TDC	Score des scores	Chiffre moyen	Valeur de F	Valeur de p	Latence	Désaturation
0	100000000	100000000	100000000	100000000	100000000	100000000
1	100000000	100000000	100000000	100000000	100000000	100000000

Fig 1 : interaction circonférence du tour de cou et index de désaturation

Graphique des interactions pour IXA0

Effet : CAT_TDC

Barres d'erreur: ± 1 Erreur(s) standard



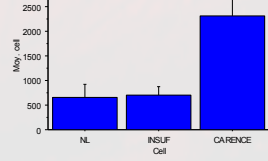
CAT_TDC	Score des scores	Chiffre moyen	Valeur de F	Valeur de p	Latence	Désaturation
0	100000000	100000000	100000000	100000000	100000000	100000000
1	100000000	100000000	100000000	100000000	100000000	100000000

Fig 2 : interaction circonférence du tour de cou et index d'apnées

Graphique des interactions pour LAT

Effet : CAT_VITD

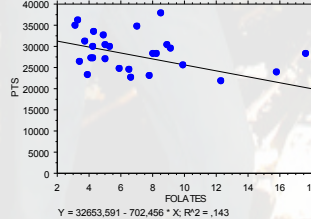
Barres d'erreur: ± 1 Erreur(s) standard



CAT_VITD	Score des scores	Chiffre moyen	Valeur de F	Valeur de p	Latence	Désaturation
0	100000000	100000000	100000000	100000000	100000000	100000000
1	100000000	100000000	100000000	100000000	100000000	100000000

Fig 3 : interaction concentration en concentration en Vitamine D et latence d'endormissement

Graphique de régression



CAT_VITD	Score des scores	Chiffre moyen	Valeur de F	Valeur de p	Latence	Désaturation
0	100000000	100000000	100000000	100000000	100000000	100000000
1	100000000	100000000	100000000	100000000	100000000	100000000

Fig 4 : interaction concentration en Folates et période totale de sommeil

Conclusion :

Les techniques exploratoires de « Data Mining » ouvrent de nouvelles voies pour l'exploration des liens pouvant exister entre le sommeil et d'autres pathologies.

Bibliographie :

- ① Laxminarayan & al ; Mining statistically significant associations for exploratory analysis of human sleep data; *IEEE Trans Inf Technol Biomed.* 2006 Jul;10(3):440-50.
- ② Zhu & al ; Introduction to medical data mining; *Sheng Wu Yi Xue Gong Cheng Xue Za Zhi.* 2003 Sep;20(3):559-62.
- ③ Jacobsen JH & al ; Factors associated with excessive daytime sleepiness in patients with severe obstructive sleep apnea ; *Sleep Breath.* 2012 Jun 24