

Comparación entre el análisis automático y manual de la polisomnografía convencional en el diagnóstico del síndrome de apnea-hipopnea obstructiva del sueño

B. Barreiro^a, G. Badosa^a, S. Quintana^b, L. Esteban^a y J.L. Heredia^a

^aServicio de Neumología. Hospital Mútua de Terrassa. Terrassa. Barcelona.

^bServicio de Cuidados Intensivos. Hospital Mútua de Terrassa. Terrassa. Barcelona. España.

OBJETIVO: Comparar el análisis automático y manual de las variables neurológicas y respiratorias obtenidas por el polisomnógrafo de 16 canales SomnoStar α 4100.

PACIENTES Y MÉTODO: Se incluyó en el estudio a 28 pacientes con sospecha de síndrome de apnea-hipopnea obstructiva del sueño a los cuales se les practicó una polisomnografía convencional. Se decidió de forma aleatoria el orden de las lecturas automática y manual de los episodios respiratorios, fases de sueño y *arousals*. Se realizó un análisis de concordancia (coeficiente de correlación intraclase), así como una representación gráfica de las diferencias utilizando el método de Bland y Altman.

RESULTADOS: Se observó una mala concordancia entre los dos tipos de análisis respecto a las fases de sueño, sobre todo REM y las fases de sueño profundo. Respecto a los parámetros respiratorios la concordancia fue buena para las apneas. Sin embargo, el análisis automático infraestimó las hipopneas. Si se considera el análisis manual como patrón de referencia para un punto de corte de índice de apneas-hipopneas mayor de 10, el análisis automático obtuvo una sensibilidad del 55%, una especificidad y un valor predictivo positivo del 100%, un valor predictivo negativo del 47% y una eficacia diagnóstica global del 67,8%.

CONCLUSIONES: El análisis automático del sistema SomnoStar 4100 proporciona una lectura inadecuada de las fases de sueño así como de los episodios respiratorios, fundamentalmente de las hipopneas.

Palabras clave: Síndrome de apnea-hipopnea obstructiva del sueño. Polisomnografía. Diagnóstico.

Comparison Between Automatic and Manual Analysis in the Diagnosis of Obstructive Sleep Apnea-Hypopnea Syndrome

OBJECTIVE: To compare automatic and manual analysis of neurological and respiratory variables obtained with the SomnoStar α 4100, a 16-channel polysomnographic system.

PATIENTS AND METHOD: Twenty-eight patients suspected of obstructive sleep apnea-hypopnea syndrome were enrolled and given conventional polysomnographic tests. The order of automatic and manual reading of respiratory episodes, sleep stages, and arousals was randomized. We assessed agreement with the intraclass correlation coefficient and plotted standardized differences against standardized means, using the Bland-Altman method.

RESULTS: Poor agreement was observed between the 2 types of analysis of sleep stages, especially for REM and deep sleep stages. Agreement was good for apneic episodes among the respiratory variables; however, automatic analysis underestimated hypopneas. If manual analysis is considered the gold standard at the apnea-hypopnea index cut point greater than 10, automatic analysis obtained a sensitivity of 55%, a specificity and positive predictive value of 100%, a negative predictive value of 47%, and an overall diagnostic yield of 67.8%.

CONCLUSIONS: The automatic analysis of the SomnoStar 4100 system provides an unsatisfactory reading of sleep stages and respiratory episodes, especially hypopneas.

Key words: Obstructive sleep apnea-hypopnea syndrome. Polysomnography. Diagnosis.

Introducción

El síndrome de apnea-hipopnea obstructiva del sueño (SAHOS) es una patología con una prevalencia que oscila entre el 1 y el 4% de la población general^{1,2}. En la actualidad la polisomnografía se considera la prueba diag-

nóstica para establecer el diagnóstico y valorar la gravedad del SAHOS. Clásicamente la lectura de las fases de sueño se realiza de forma manual según los criterios establecidos previamente³. Sin embargo, existe variabilidad interindividual del análisis polisomnográfico y además requiere mucho tiempo y recursos. Los polígrafos modernos incorporan sistemas de análisis automático de los parámetros neurológicos y, asimismo, determinan de forma automática los episodios respiratorios, las desaturaciones de la oxihemoglobina y los movimientos respiratorios. Dichos sistemas automáticos no están suficientemente validados y en la práctica clínica son poco

Correspondencia: Dr. B. Barreiro López.
Servicio de Neumología. Hospital Mútua de Terrassa.
Pza. Dr. Robert, 5. 08221 Terrassa. Barcelona. España.
Correo electrónico: pneumologia@mutuaterrassa.es

Recibido: 5-3-2003; aceptado para su publicación: 1-7-2003.

precisos en la estadiación de las fases de sueño y en la detección de los episodios respiratorios. Dadas las diferencias existentes entre los diversos tipos de análisis, se planteó un estudio con el objetivo de comparar el análisis automático y manual de los parámetros obtenidos por el polisomnógrafo de 16 canales Somnostar α 4100 (SensorMedics Corporation, California, EE.UU.).

Pacientes y método

El ámbito de estudio ha sido el Hospital Mútua de Terrassa, hospital de nivel II situado en Terrassa, con una cobertura asistencial de 200.000 habitantes. Dicho hospital dispone de una Unidad de Sueño adscrita al Servicio de Neumología con capacidad para realizar polisomnografía convencional y polisomnografía respiratoria.

Durante un período de tres meses se estudiaron 28 pacientes remitidos desde la consulta externa de Neumología con el diagnóstico de sospecha de SAHOS. A todos se les realizaron radiografía de tórax, espirometría forzada, analítica y cuestionario de Epworth. También se les realizó una polisomnografía convencional (Somnostar 4100) vigilada en la Unidad de Sueño de nuestro centro. Se monitorizaron los siguientes parámetros: 4 derivaciones de electroencefalograma (EEG; C4-A1, C3-A2, O1-A2, O2-A1), electrooculograma, electromiograma mentoniano, electromiograma tibial, flujo nasobucal mediante termistor, movimientos toracoabdominales mediante sensores piezoeléctricos, electrocardiograma y saturación arterial mediante pulsioxímetro. No se monitorizó la onda de presión nasal por no disponer del equipo, lo que constituye una limitación de estudio. Se definió la apnea como cese del flujo oronasal de 10 s de duración, y la hipopnea como la reducción significativa del flujo oronasal y/o movimientos toracoabdominales acompañada de *arousals* y/o una desaturación cíclica del 3% o superior. Se definió el *arousal* como un aumento de la frecuencia del EEG de más de 3 s de duración sujeta a varias condiciones, siguiendo la normativa de la American Sleep Disorders Association⁴. El diagnóstico de SAHOS se estableció a partir de un índice apnea-hipopnea (IAH) mayor de 10/h obtenido por polisomnografía convencional. Ninguno de los pacientes había iniciado previamente tratamiento con presión positiva continua de la vía aérea. Uno de los investigadores (B.B.) efectuó una lectura manual y automática de la polisomnografía variando aleatoriamente el orden de realización. La lectura automática del Somnostar 4100 marca el trazado de la polisomnografía. Dichas marcas se eliminaron para realizar la lectura manual y, por tanto, no interfirieron en la estadiación posterior. La estadiación manual de las fases de sueño se realizó según los parámetros previamente establecidos por Rechtschaffen y Kales³. La lectura automática del EEG se efectuó mediante el *software* del Somnostar 4100, que utiliza un análisis espectral. Un algoritmo matemático identifica la amplitud y la frecuencia de las ondas del EEG y los clasifica en delta, theta, alpha y beta. Dicho algoritmo también se aplica a la señal de electrooculograma. Los episodios respiratorios se determinaron de forma automática mediante el sistema de análisis del Somnostar 4100. Dicho sistema determina la línea de base promediando el número de respiraciones en los 2 min previos a la aparición del episodio. Define la apnea como una reducción del flujo oronasal mayor del 80% respecto a la línea de base, y la hipopnea como una caída del flujo oronasal de un 50% respecto a la línea de base más un 4% de desaturación. Los resultados se expresan en forma de media \pm desviación estándar. Para conocer la concordancia entre los dos tipos de análisis se utilizó el coeficiente de correlación intraclase. Para representar gráficamente las

TABLA I
Características antropométricas y espirométricas de los 28 pacientes con sospecha de síndrome de apnea-hipopnea obstructiva del sueño

Edad (años)	50 \pm 8,8
FVC (ml)	3.100 \pm 1.300
FVC (%)	85 \pm 28
FEV ₁ (ml)	2.900 \pm 1.300
FEV ₁ (%)	80 \pm 27
Escala Epworth	14 \pm 5
IMC (kg/m ²)	33 \pm 5,2

Los datos se presentan como media \pm desviación estándar. FVC: capacidad vital forzada; FEV₁: volumen espiratorio forzado en el primer segundo; IMC: índice de masa corporal.

TABLA II
Concordancia entre los análisis manual y automático de la polisomnografía

	CCI	IC del 95%	p
Fase 1	0,551	0,23-0,76	0,0009
Fase 2	0,518	0,18-0,74	0,002
Fase 3	0,057	-0,31-0,41	0,383
REM	0,450	0,099-0,70	0,0071
IAH	0,807	0,62-0,90	0,00001
Índice hipopneas/h	0,276	-0,10-0,58	0,07
Índice apneas/h	0,979	0,95-0,99	0,0001

CCI: coeficiente de concordancia intraclase; IC: intervalo de confianza; IAH: índice de apneas-hipopneas. p < 0,05, diferencia significativa.

diferencias entre ambos tipos de análisis se empleó el método de Bland y Altman⁵, dado que lo que se pretende es buscar concordancia entre dos variables cuantitativas que deberían dar el mismo resultado con dos métodos distintos. Se calcularon la sensibilidad, la especificidad, los valores predictivos positivo y negativo de los parámetros respiratorios obtenidos mediante el análisis manual teniendo como patrón de referencia un IAH por polisomnografía convencional superior a 10/h. Se consideró estadísticamente significativo un valor de p < 0,05.

Resultados

Se estudiaron 28 pacientes con una edad media de 50 años, con una distribución por sexos de 21 varones y 7 mujeres. Las características antropométricas y funcionales quedan reflejadas en la tabla I. Cabe destacar que eran pacientes con obesidad moderada y con marcada somnolencia diurna. El diagnóstico final establecido mediante análisis manual fue: 20 casos de SAHS y 8 no SAHS. Se observó una moderada concordancia entre el análisis manual y automático respecto a los parámetros de sueño, así como en la mayoría de los parámetros respiratorios (tabla II). El análisis automático tiende a infraestimar el sueño REM (p < 0,007) y las fases de sueño profundo (p < 0,3). Sin embargo, dicha concordancia es moderada respecto al sueño superficial (fases 1 y 2). En lo que respecta a los parámetros respiratorios, la concordancia entre ambos análisis es buena para el IAH final (p < 0,0001) así como las apneas (p < 0,0001). Sin embargo, la concordancia respecto a las hipopneas es mala, ya que el análisis automático las infraestima. Desde el punto de vista gráfico existen grandes diferencias entre los dos análisis respecto a las fases de sueño debi-

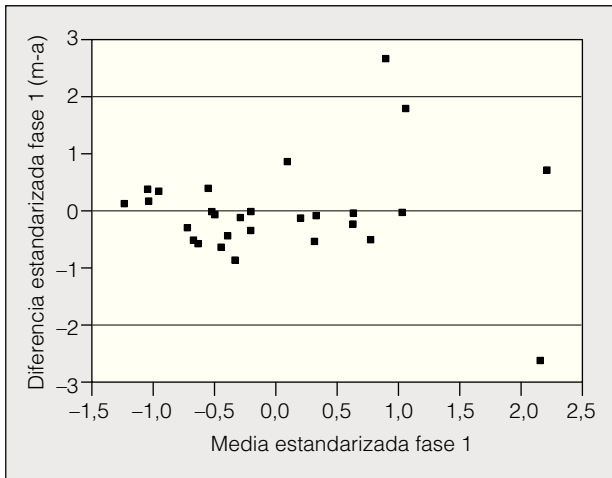


Fig. 1. Comparación de la diferencia estandarizada entre los análisis manual (m) y automático (a) de la fase 1 con la media. Las líneas horizontales representan los límites superior e inferior del 95% del intervalo de confianza.

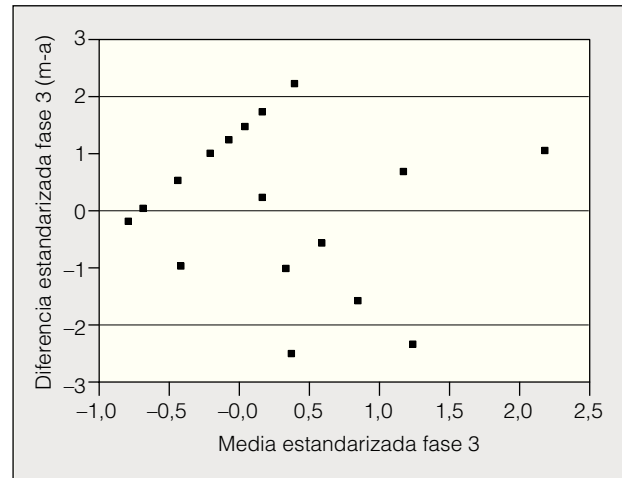


Fig. 2. Comparación de la diferencia estandarizada entre los análisis manual (m) y automático (a) de la fase 3 con la media. Las líneas horizontales representan los límites superior e inferior del 95% del intervalo de confianza.

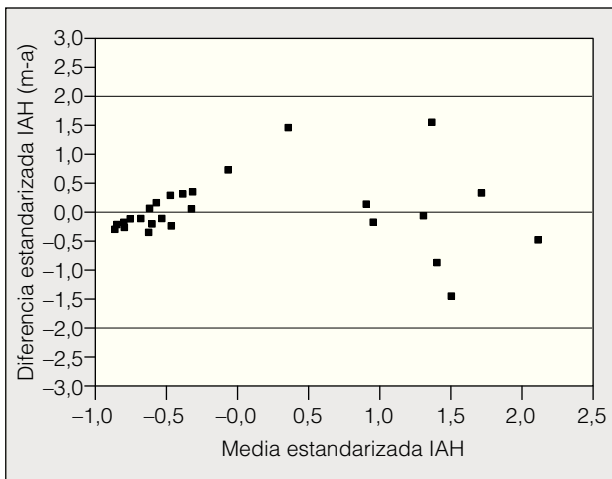


Fig. 3. Comparación de la diferencia estandarizada del índice de apneas-hipopneas (IAH) entre los análisis manual (m) y automático (a) con la media. Las líneas horizontales representan los límites superior e inferior del 95% del intervalo de confianza.

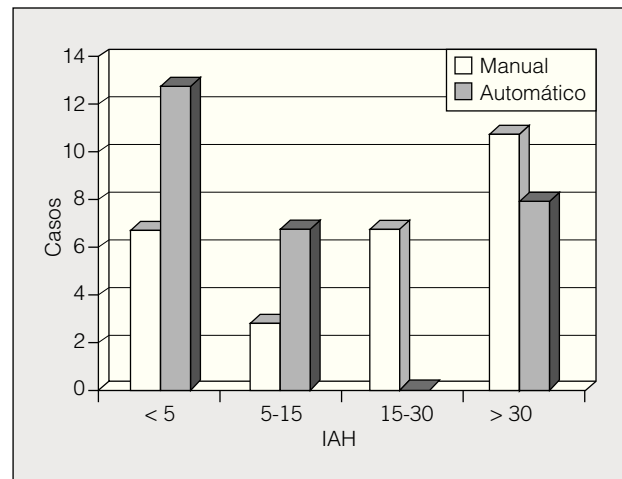


Fig. 4. Estratificación de los episodios respiratorios en los análisis automático y manual. IAH: índice de apneas-hipopneas.

do, fundamentalmente, a la escasa precisión del análisis automático (figs. 1 y 2). La comparación de los episodios respiratorios muestra escasas diferencias respecto al IAH (fig. 3). Sin embargo, a medida que aumenta el número de episodios, fundamentalmente hipopneas, la concordancia entre los métodos es claramente inferior.

Al realizar un análisis estratificado de los episodios respiratorios dentro del grupo de pacientes con un IAH superior a 30, el análisis manual aporta pocos diagnósticos nuevos. Sin embargo, si el IAH está entre 15 y 30, el análisis manual ofrece un mayor número de diagnósticos en 7 de los 28 casos (25%; fig. 4).

Si consideramos el análisis manual como patrón de referencia, el análisis automático para un punto de corte de IAH > 10 obtiene una sensibilidad del 55%, una especificidad del 100%, un valor predictivo del 100%, un valor predictivo negativo del 47% y una eficacia diagnóstica global del 67,8%.

Discusión

El presente estudio confirma que el análisis automático de los parámetros neurológicos y respiratorios del polisomnógrafo Somnostar 4100 es menos sensible que el análisis manual. Existe una buena concordancia entre los dos tipos de análisis respecto al IAH y una mala concordancia respecto a las fases de sueño, especialmente las de sueño profundo y REM.

Los métodos de análisis automático de los parámetros respiratorios pueden ser de ayuda ya que proporcionan información sobre parámetros adicionales como la duración de los episodios, la saturación media y mínima, el CT-90 (porcentaje del tiempo de registro con saturación < 90%), cuantificación del ronquido y posición corporal. Dichos métodos tienden a infraestimar el IAH con respecto al análisis manual, debido fundamentalmente a que no se reconocen las hipopneas⁶. La sensibi-

lidad y especificidad del análisis automático son variables según los criterios considerados. En nuestro estudio el análisis automático infraestima el IAH, especialmente si el número de episodios respiratorios es bajo (< 30/h) y si predominan las hipopneas. Asimismo, considerando un IAH mayor de 10, la sensibilidad y el valor predictivo negativo del análisis automático fueron del 55 y el 47%, respectivamente. Esto probablemente esté ligado a la falta de detección de las hipopneas, y por ello es necesario un análisis manual de los parámetros respiratorios. Se han comunicado resultados similares en el estudio de Zucconi et al⁷, donde el análisis automático y/o semiautomático de los parámetros respiratorios tiene una sensibilidad y especificidad altas para IAH elevados, no así para IAH bajos. Sin embargo, algunos autores⁸ encuentran una buena correlación entre los dos análisis respecto al IAH. Ello depende en gran medida de los sistemas automáticos utilizados. Se han valorado sistemas simplificados de análisis comparados con polisomnografía convencional⁹. En estos sistemas el análisis manual asistido no mejora la rentabilidad diagnóstica del análisis automático. Otros autores^{10,11} confirman que el análisis manual es superior al automático.

Los sistemas automáticos de análisis del sueño han mejorado en los últimos años. Sin embargo, dichos sistemas infraestiman el tiempo total de sueño, la fase 2 debido fundamentalmente a una dificultad en la identificación de las ondas K y los *spindles*. Asimismo, sobrestiman la fase 1 y afectan en poca medida a las fases 3 y REM¹². En el presente estudio la concordancia entre los dos tipos de análisis es moderada respecto a las fases de sueño superficial y mala respecto a las fases de sueño profundo y REM.

Existen diferentes sistemas de análisis del EEG utilizando el índice de análisis espectral¹³. La principal ventaja de este método comparado con el análisis visual es que las fases de sueño profundo son valoradas más objetivamente y de forma continua.

Existen métodos de detección automática de *spindles*¹⁴. Es un método computarizado que permite un cálculo cuantitativo de la frecuencia y amplitud de las ondas EEG y *spindles*. Además, este tipo de análisis permite reducir el número de artefactos y, por tanto, es un método muy flexible.

Philip-Joet et al¹⁵, utilizando un análisis espectral de EEG, consiguen una concordancia total con el análisis manual del 81%, una concordancia parcial del 11% y discordancia en el 8%. Este sistema permite una rápida evaluación de la calidad del EEG. Sin embargo, en el presente estudio la concordancia entre los dos tipos de análisis respecto a las fases de sueño es mala, especialmente en lo que se refiere a las fases de sueño profundo y REM. Probablemente el programa de análisis automático no identifica de forma correcta los *spindles* y las ondas K. Asimismo no identifica correctamente la fase REM, que en ocasiones confunde con la fase 1, valorando incorrectamente los movimientos oculares.

Hay varios factores que pueden modificar las características y la interpretación del EEG. En primer lugar, el llamado efecto de la primera noche produce un aumento del período de despierto, una disminución del tiempo

total de sueño, una disminución de la eficiencia de sueño y una disminución de la fase REM¹⁶. En segundo lugar, otro factor que influye es la variabilidad interobservador, con un grado de concordancia entre los diversos técnicos que oscila entre el 82 y el 88%^{17,18}. En el presente estudio dicho factor no se ha tenido en cuenta ya que las lecturas las realizó el mismo investigador. En tercer lugar existe una variabilidad intraobservador que no se ha evaluado en el presente estudio. Dicho factor puede modificar ligeramente el análisis manual de la polisomnografía y por ello limitar los resultados.

En la actualidad los sistemas de análisis automático de los polisomnógrafos tienen una sensibilidad y especificidad limitadas⁶, ya que proporcionan una lectura inadecuada de algunos episodios respiratorios (hipopneas) y de las fases de sueño. No obstante, el análisis automático puede simplificar las lecturas y actualmente se recomienda realizar en primer lugar un análisis automático de la polisomnografía y posteriormente una revisión manual¹⁹.

En resumen, el análisis manual de la polisomnografía convencional es el método más sensible y específico para la correcta estratificación de las fases de sueño y de los episodios respiratorios. Es importante validar nuevos sistemas automáticos de análisis para su utilización en la práctica clínica diaria, aumentando de este modo los recursos disponibles.

Agradecimientos

Los autores desean agradecer la colaboración del Dr. F. Barbé, del Hospital Universitari Son Dureta, en la elaboración del presente manuscrito.

BIBLIOGRAFÍA

1. Young T, Palta M, Dempsey J, Skatrud J, Weber S, Badr S. The occurrence of sleep-disordered breathing among middle-aged adults. *N Engl J Med* 1993;318:1230-5.
2. Durán J, Esnaola S, Ramón R, Iztueta A. Obstructive sleep apnea-hypopnea and related clinical features in a population-based sample of subjects aged 30 to 70 years. *Am J Respir Crit Care Med* 2001;163:685-9.
3. Rechtschaffen A, Kales A. A manual of standardized terminology techniques and scoring system for sleep stages of human subjects. Washington DC: Public Health Service, US Government Printing Office, 1963.
4. American Sleep Disorders Association. EEG arousals: scoring rules and examples. A preliminary report from the Sleep Disorders Association. *Sleep* 1992;15:17-184.
5. Bland JM, Altman DG. Statistical method for assessing agreement between two methods of clinical measurement. *Lancet* 1986;1:307-10.
6. Carrasco O, Monserrat JM, Lloberes P, Ascaso C, Ballester E, Fornes C, et al. Visual and different automatic scoring profiles of respiratory variables in the diagnosis of sleep apnoea-hypopnea syndrome. *Eur Respir J* 1996;9:125-30.
7. Zucconi M, Ferini-Stambi L, Castronovo V, Oldani A, Smirne S. An unattended device for sleep-related breathing disorders: validation study in suspected obstructive sleep apnoea syndrome. *Eur Respir J* 1996;9:1251-6.
8. Verse T, Pirsing W, Kroker B, Junge-Hulsing B, Zimmerman E. Validating a 7-channel ambulatory polygraphy unit: operating instructions for the physician and patient. *HNO* 1999;47:249-55.

BARREIRO B, ET AL. COMPARACIÓN ENTRE EL ANÁLISIS AUTOMÁTICO Y MANUAL DE LA POLISOMNOGRAFÍA CONVENCIONAL EN EL DIAGNÓSTICO DEL SÍNDROME DE APNEA-HIPOPNEA OBSTRUCTIVA DEL SUEÑO

9. Jiménez Gómez A, Golpe Gómez R, Carpizo Alfayete R, De la Roza Fernández C, Fernández Rozas S, García Pérez MM. Validación de un polígrafo respiratorio de 3 canales (Edentec) para el diagnóstico del síndrome de apnea del sueño. *Arch Bronconeumol* 2000;36:7-12.
10. Esnaola S, Durán J, Infante-Rivard C, Rubio R, Fernández A. Diagnostic accuracy of a portable recording device (MESAM IV) in suspected obstructive sleep apnea. *Eur Respir J* 1996;9:2597-605.
11. Koziej M, Cieslicki JK, Gorzelak K, Sliwinski P, Zielinski J. Hand-scoring of MESAM 4 recordings is more accurate than automatic analysis in screening for obstructive sleep apnoea. *Eur Respir J* 1994;7:1771-5.
12. Sforza E, Vandi S. Automatic Oxford-Medilog 9200 sleep staging scoring: comparison with visual analysis. *J Clin Neurophysiol* 1996; 13:227-33.
13. Hammer N, Todorova A, Hofman HC, Schober F, Vonderheid-Guth B, Dimpfel W. Description of healthy and disturbed sleep by means of the spectral frequency index (SFX). A retrospective analysis. *Eur J Med Res* 2001;6:333-44.
14. Schinicek P, Zeitlhofer J, Anderer P, Saletu B. Automatic sleep-spindle detection procedure: aspects of reliability and validity. *Clin Electroencephalogr* 1994;25:26-9.
15. Philip-Joet FF, Rey MF, Dicroco AA, Reynaud-Gaubert MJ, Arnaud AG. Semi-automatic analysis of electroencephalogram in sleep apnea syndromes. *Chest* 1993;104:336-9.
16. Toussaint M, Luthringer R, Schaltenbrand N, Nicolas A, Jacqmin A, Carelli G, et al. Changes in EEG power density during sleep laboratory adaptation. *Sleep* 1997;20:1201-7.
17. Scaltenbrand N, Lengelle R, Toussaint M, Luthringer R, Carelli G, Jacqmin A, et al. Sleep stage scoring using the neural network model: comparison between visual and automatic analysis in normal subjects and patients. *Sleep* 1996;19:26-35.
18. Hoelscher TJ, McCall WV, Powell J, Marsh GR, Erwin CW. Two methods of scoring sleep with the Oxford Medilog 9000: comparison to conventional paper scoring. *Sleep* 1989;12:133-9.
19. De Carli F, Nobili L, Gelcich P, Ferrillo F. A method for the automatic detection of arousals during sleep. *Sleep* 1999;22:561-72.