

# Evaluación de un protocolo de *screening* de apnea del sueño basado en el análisis automático de la oximetría no supervisada

D. Álvarez González<sup>1,2</sup>, G. C. Gutiérrez-Tobal<sup>1</sup>, A. Crespo Sedano<sup>2</sup>, C. A. Arroyo Domingo<sup>2</sup>, F. Del Campo Matías<sup>1,2</sup>, R. Hornero Sánchez<sup>1</sup>

<sup>1</sup> Grupo de Ingeniería Biomédica, Universidad de Valladolid, Valladolid, España, daniel.alvarez@gib.tel.uva.es

<sup>2</sup> Servicio de Neumología, Hospital Universitario Río Hortega, Valladolid, España, fsas@tel.uva.es

## Resumen

*El síndrome de apnea-hipopnea del sueño (SAHS) es un trastorno respiratorio relacionado con el sueño de elevada prevalencia. Si no se diagnostica correctamente puede tener graves consecuencias en la calidad de vida y en la salud de los pacientes. En el presente estudio se ha diseñado y evaluado un nuevo protocolo de screening de pacientes basado en el procesado automático de la señal de saturación de oxígeno en sangre (SpO<sub>2</sub>), registrada de forma no supervisada en el domicilio del paciente. La metodología propuesta se divide en 2 etapas complementarias: (i) etapa de clasificación, que emplea un modelo de regresión logística (LR) optimizado para detectar sujetos sin SAHS de tal forma que no se produzcan falsos negativos y (ii) etapa de estimación del índice de apnea hipopnea (IAH), que emplea una red neuronal perceptrón multicapa (MLP) entrenada para estimar el IAH y detectar a los pacientes SAHS positivo más severos sin que se produzcan falsos positivos. Para entrenar y validar la metodología se ha empleado una población de 299 pacientes consecutivos dividida en entrenamiento (50%) y test (50%). La metodología de screening de 2 fases propuesta evitaría la realización del 40.7% de las PSGs completas en el hospital, sin que se produzcan falsos negativos (valor predictivo negativo, NPV = 100%) ni falsos positivos (valor predictivo positivo, PPV = 100%).*

## 1. Introducción

El síndrome de la apnea-hipopnea del sueño (SAHS) es una enfermedad crónica que se caracteriza por la aparición recurrente de eventos de apnea (cese total de la respiración) e hipopnea (reducción significativa del flujo aéreo) durante el sueño [1]. La repetición de estos eventos a lo largo de la noche ocasiona despertares transitorios que impiden que el sueño sea reparador [2], provocando graves consecuencias para la salud y calidad de vida de los pacientes [1]. Por este motivo, un diagnóstico temprano resulta fundamental para controlar su desarrollo y aplicar el tratamiento apropiado. Sin embargo, la prueba estándar de diagnóstico, la polisomnografía nocturna (PSG), presenta importantes limitaciones que dificultan tanto la rapidez para obtener un diagnóstico como el correspondiente tratamiento [3]. Por ello, la búsqueda de alternativas eficaces que permitan simplificar el diagnóstico y reducir las listas de espera se ha convertido en un reto de gran interés científico, médico y económico.

Las principales limitaciones de la PSG radican en su complejidad técnica. Durante la prueba se requiere la monitorización de hasta 32 señales fisiológicas de los

pacientes mientras duermen, lo que requiere costosos equipos de adquisición de señales y personal de supervisión [4]. Además, la PSG ha de realizarse en una unidad del sueño especializada, fuera del entorno de sueño habitual de los pacientes. Los problemas de adaptación a este entorno pueden condicionar la prueba, pudiendo hacer necesaria su repetición [3].

El enfoque planteado en este estudio se basa en la simplificación de la prueba diagnóstica mediante el análisis automático de una única señal: la saturación de oxígeno en sangre (SpO<sub>2</sub>) procedente de la pulsioximetría nocturna. En los últimos años, numerosos estudios han demostrado que su empleo de manera individual puede resultar de gran utilidad en la ayuda al diagnóstico del SAHS [5, 6]. Además, se adquiere de manera sencilla mediante un pulsioxímetro colocado en el dedo del paciente. Esta sencillez ha derivado en la existencia de pequeños dispositivos portátiles comerciales que facilitan la realización de la prueba en el domicilio del paciente de forma no supervisada.

El presente estudio se plantea bajo la hipótesis de que la aplicación de técnicas de procesado avanzado permitirán maximizar el potencial diagnóstico y extraer toda la información útil de la señal de SpO<sub>2</sub>. El objetivo principal del estudio consiste en reducir en un porcentaje significativo el número de PSGs completas necesarias sin que se produzcan errores diagnósticos (falsos positivos o falsos negativos). Para ello, se diseñará y evaluará un nuevo protocolo de *screening* de pacientes sospechosos de padecer SAHS basado en el análisis automático de la señal de SpO<sub>2</sub> obtenida en el domicilio del paciente. Este estudio se aborda desde la perspectiva del reconocimiento de patrones, que permite afrontar el diagnóstico automático mediante dos enfoques distintos [7]: la clasificación binaria de pacientes en “SAHS” o “no SAHS” y la regresión (estimación) del índice de apnea-hipopnea (IAH), que permite caracterizar la severidad de la enfermedad. Para ello, se implementarán las siguientes etapas: (i) clasificación binaria basada en un modelo multivariante de regresión logística (LR), optimizado para identificar a pacientes que no padecen SAHS con un valor predictivo negativo del 100% y (ii) estimación del IAH mediante una red neuronal perceptrón multicapa (MLP) de regresión multivariante, optimizada para detectar con toda seguridad pacientes de SAHS con un valor predictivo positivo del 100%.

## 2. Base de datos de estudios de sueño

La población bajo estudio estaba compuesta por 299 sujetos sospechosos de padecer SAHS derivados consecutivamente a la unidad del sueño del Hospital Universitario Río Hortega de Valladolid entre el 10/06/2013 y el 27/03/2014. Cada sujeto fue informado y preguntado sobre su participación en el estudio, dando su consentimiento por escrito. La población se dividió en grupo de entrenamiento (149 primeros sujetos, 50%), empleado para encontrar la configuración óptima de la metodología de *screening* propuesta, y grupo de test (150 sujetos siguientes, 50%) empleado para validar el protocolo. La Tabla 1 muestra las características sociodemográficas y clínicas de la población bajo estudio.

Cada sujeto realizó una PSG completa en el hospital (polisomnógrafo E-Series, Compumedics). Un único especialista calculó el IAH derivado de la PSG para determinar el diagnóstico y severidad de la patología en base a los criterios de la *American Academy of Sleep Medicine*. Se diagnosticó SAHS a los pacientes con un IAH igual o superior a 10 eventos por hora (e/h) de sueño. Cada paciente realizó la prueba de oximetría domiciliaria no supervisada en la noche inmediatamente anterior o posterior (asignación aleatoria) a la PSG (Nonin WristOx2 3150, Nonin Medical Inc). La frecuencia de muestreo de la señal de SpO<sub>2</sub> fue de 1 Hz. Los registros de oximetría fueron exportados desde el dispositivo portátil a un ordenador para su posterior análisis *offline*.

## 3. Metodología

Nuestra propuesta plantea la realización de un *screening* automático en la consulta de Atención Primaria. En la Figura 1 se muestra un diagrama de bloques que ilustra el protocolo propuesto. Los métodos de procesado automático de la señal de SpO<sub>2</sub> comprenden tres etapas: (i) extracción de características, (ii) selección de características y (iii) reconocimiento de patrones.

### 3.1. Extracción de características

Se derivaron características de la señal de SpO<sub>2</sub> de diferente naturaleza para obtener la mayor cantidad posible de información útil y complementaria. El conjunto inicial de características estaba formado por las siguientes variables:

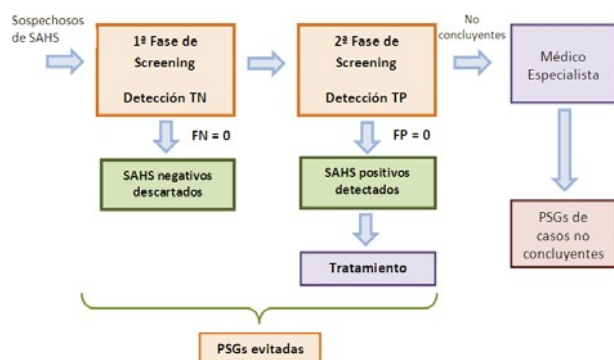


Figura 1. Diagrama de bloques del protocolo de screening

Todos	Entrenamiento	Test	
Sujetos (n)	299	149	150
Edad (años)	55.12 ± 13.33	55.27 ± 12.55	54.97 ± 14.10
Varones (n)	226 (75.6%)	117 (78.5%)	109 (72.7%)
IMC (kg/m <sup>2</sup> )	29.48 ± 5.59	29.63 ± 5.23	29.34 ± 5.93
Registro (h)	7.41 ± 1.54	7.33 ± 1.41	7.48 ± 1.67
IAH (e/h)	41.55 ± 28.99	42.63 ± 28.16	40.48 ± 29.85

Tabla 1. Características socio-demográficas y clínicas

- Parámetros espectrales. Los recurrentes eventos de apnea provocan cambios en el espectro de los registros de SpO<sub>2</sub>. Se aplicó una ventana de Hanning ( $L = 512$ ), con solapamiento del 50% y DFTs de 1024 puntos para estimar la PSD mediante el método de Welch [8, 9]. El contenido espectral se parametrizó mediante 5 características: (i) potencia total ( $P_T$ ), (ii) amplitud de pico ( $PA$ ) en la banda [0.014 – 0.033 Hz], (iii) potencia relativa ( $P_R$ ) en la banda [0.014 – 0.033 Hz], (iv) frecuencia mediana ( $FM$ ) y (v) entropía espectral de Shannon ( $SE$ ).
- Parámetros no lineales. Se aplicaron los siguientes métodos para caracterizar la naturaleza no lineal de los cambios sufridos por los registros de SpO<sub>2</sub> [8]: (i) entropía muestral ( $SampEn$ :  $m=1$ ,  $r=0.25*SD$ ), que cuantifica la irregularidad; (ii) medida de la tendencia central ( $CTM$ :  $r=1$ ), que estima la variabilidad; y (iii) complejidad de Lempel-Ziv ( $LZC$ ), que cuantifica la complejidad. Se fijó un tamaño de ventana móvil de  $L=512$  muestras sin solapamiento, tomando el promedio a lo largo del tiempo como valor final de cada parámetro.

### 3.2. Selección de características

Se aplicó el método de selección bidireccional paso a paso hacia delante (FSFS) propuesto por Hosmer & Lemeshow [10]. Esta metodología de exploración da lugar a un subconjunto anidado de características, como resultado de ir incorporando secuencialmente las variables más relevantes y eliminando aquellas que aportan información redundante a un modelo LR en cada paso del algoritmo [11]. Las principales ventajas de los métodos secuenciales son la eficiencia computacional y la robustez frente al sobre-entrenamiento [11].

### 3.3. Reconocimiento de patrones

Finalmente, se desarrollaron métodos de reconocimiento de patrones desde dos enfoques diferentes: clasificación binaria (identificación de sujetos “SAHS negativo” y “SAHS positivo”) y regresión (estimación del IAH).

- Clasificación. Un clasificador LR binario permite estimar la probabilidad de que se produzca el suceso definido por una variable dependiente dicotómica en función de una o varias variables explicativas o independientes [12]. LR modela la función de densidad de probabilidad como una distribución de Bernuilli y emplea la razón de máxima verosimilitud para determinar los coeficientes del modelo de clasificación [12].

- Regresión. Los clasificadores binarios no proporcionan información sobre el grado de severidad, lo que resultaría especialmente útil en el grupo SAHS positivo, en el que la patología puede ser leve, moderada o severa en función del IAH. Las redes neuronales perceptrón multicapa (MLP) tienen un gran potencial para el modelado de problemas de regresión gracias a su arquitectura flexible [7]. El parámetro de regularización ( $\alpha$ ) y número de neuronas de la capa oculta ( $N_H$ ) se optimizaron en términos del índice de correlación intraclase (ICC) entre el IAH estimado y el IAH de la PSG aplicando validación cruzada dejando uno fuera (loo-cv) en entrenamiento. Se aplicó una función lineal de activación y se empleó una única neurona de salida para modelar la variable objetivo (IAH).

### 3.4. Análisis estadístico

El análisis del rendimiento diagnóstico se realizó en base a las siguientes métricas: índice de correlación intraclase (ICC) verdaderos positivos (TP), verdaderos negativos (TN), falsos positivos (FP), falsos negativos (FN), sensibilidad (Se), especificidad (Sp), valor predictivo positivo (PPV), valor predictivo negativo (NPV), precisión (Ac) y área bajo la curva ROC (AROC).

## 4. Resultados

### 4.1. Entrenamiento

El algoritmo FSFS seleccionó 5 características: 2 lineales en el dominio de la frecuencia ( $PA$  y  $PR$ ) y 3 no lineales en el dominio del tiempo ( $SampEn$ ,  $CTM$  y  $LZC$ ). La Tabla 2 muestra el rendimiento diagnóstico del modelo LR construido con estas características empleando 2 umbrales: (i) par Se vs. Sp balanceado ( $u_{opt} = 0.83$ ) y (ii) FN igual a cero ( $u_{oFN} = 0.22$ ).

La red MLP se optimizó ( $\alpha=1$  y  $N_H=11$ ) y reentrenó en el conjunto de entrenamiento, alcanzando un ICC=0.800. La Figura 2 muestra la correlación entre el IAH real y el IAH estimado en entrenamiento. Es posible observar un sesgo, que ha sido estimado mediante ajuste por mínimos cuadrados ( $bias=14.0219$ ) y eliminado a la salida. La Tabla 2 muestra el rendimiento diagnóstico de la red empleando 2 umbrales: (i) par Se vs. Sp balanceado ( $u_{opt} = 10.36$  e/h) y (ii) FP igual a cero ( $u_{oFP} = 30$  e/h).

### 4.2. Test

La Figura 3 muestra la distribución de datos a la salida del clasificador LR para la población de test ( $N=150$ ). Se obtuvieron 4 TN con NPV=100.0%, lo que implica que el modelo clasifica sujetos sin SAHS con total certeza (Se: 100.0%; Sp: 18.2%; Ac: 88.0). La Figura 4 muestra la correlación entre el IAH real y el IAH estimado por la red MLP (ICC = 0.650) aplicada a los sujetos procedentes de la primera fase ( $N=146$ ). Se obtuvieron 57 TP con PPV=100.0%, lo que implica que la red clasifica sujetos con SAHS con total certeza (Se: 44.5%; Sp: 100.0%; Ac: 51.4). Aplicando el protocolo de doble *screening* propuesto en el presente estudio sería posible evitar la realización de 61 PSGs (40.7%) completas en el hospital.

Modelo	TP	TN	FP	FN	Se	Sp	PPV	NPV	Ac	AROC
LR $u_{opt}$	110	18	2	19	85.3	90.0	98.2	48.7	85.9	0.94
LR $u_{oFN}$	129	2	18	0	100.0	10.0	87.8	100.0	87.9	0.94
MLP $u_{opt}$	109	19	1	20	84.5	95.0	99.1	48.7	85.9	0.93
MLP $u_{oFP}$	54	20	0	75	41.9	100.0	100.0	21.1	49.7	0.93

Tabla 2. Rendimiento diagnóstico individual del modelo LR y de la red MLP en el conjunto de entrenamiento ( $N=149$ )

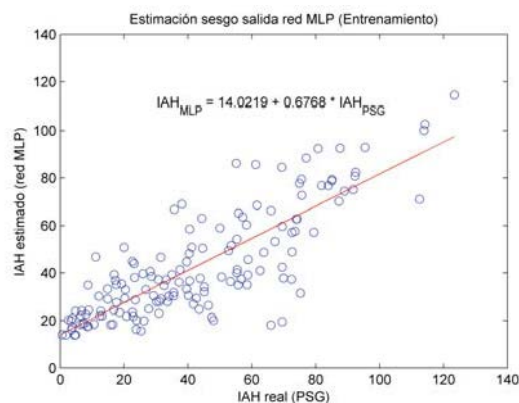


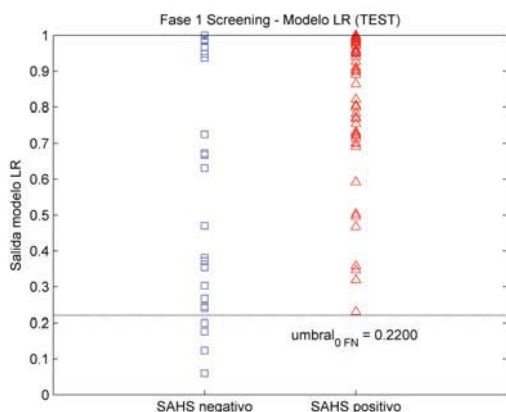
Figura 2. Estimación del sesgo a la salida de la red MLP

## 5. Discusión y conclusiones

En el presente estudio se ha evaluado el rendimiento de un nuevo protocolo de doble *screening* basado en el procesamiento automático del registro portátil de  $SpO_2$  realizado en el domicilio del paciente para la ayuda al diagnóstico del SAHS. Para ello, se ha diseñado y validado una metodología de despistaje compuesta por 2 fases complementarias: (i) clasificador LR, entrenado para detectar sujetos sin SAHS sin que se produzcan falsos negativos (FN=0, NPV=100.0%) y (ii) red neuronal MLP de regresión, entrenada para detectar la patología sin que se produzcan falsos positivos (FP=0, PPV=100.0%). Los umbrales conservadores determinados en el conjunto de entrenamiento mantuvieron la misma eficacia en el conjunto de test, permitiendo detectar 61 casos (4 SAHS negativo y 57 SAHS positivo) en los que sería posible evitar la realización de una PSG completa.

Los resultados obtenidos concuerdan con los alcanzados en estudios previos que evalúan diferentes metodologías de *screening* de SAHS en múltiples contextos. En el estudio desarrollado por Gurubhagavatula *et al* se evaluó una estrategia de *screening* de SAHS severo (IAH2:30) en pacientes hipertensos basada en la combinación de parámetros clínicos y el IAH derivado de la poligrafía simplificada no supervisada en el domicilio [13]. Esta metodología alcanzó valores de Se variando entre 62.9% y 94.2%, Sp entre 68.9% y 84.0% y NPV entre 88.1% y 98.5%, en el grupo de validación. En un estudio reciente desarrollado por Torres *et al* se combinaron estadísticos derivados de la monitorización ambulatoria de la presión sanguínea, variables antropométricas, sociodemográficas y biológicas mediante un modelo LR para discriminar pacientes con SAHS moderado a severo (IAH2:15) en pacientes con hipertensión [14]. Esta metodología alcanzó un AROC de 0.80 en entrenamiento y un par Se vs. Sp de 100.0% y 66.7% en el grupo de validación, respectivamente. Crowley *et al* evaluaron la capacidad





**Figura 3.** Distribución de valores a la salida del modelo LR de clasificación aplicado sobre la población de test

diagnóstica de la señal de flujo aéreo de sonda de presión registrada en el domicilio mediante el dispositivo portátil *ApneaLink* [15]. El IAH estimado por el dispositivo alcanzó un par Se vs. Sp de 81% vs. 77% y 67% vs. 91% para puntos de corte de IAH = 5 y 15, respectivamente.

Es necesario tener presentes varias limitaciones. Es importante destacar que la PSG de referencia se realizó en una noche diferente a la prueba domiciliaria, con lo que la variabilidad entre noches del IAH podría influir en los resultados. Además, el número de sujetos SAHS negativo es bastante reducido, por lo que una muestra más balanceada proporcionaría resultados más generalizables.

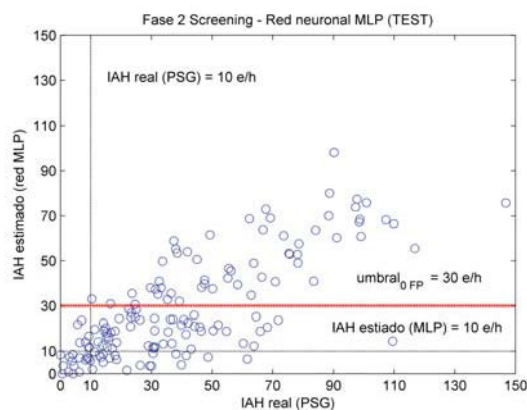
Los resultados sugieren que los dispositivos de pulsioximetría portátiles podrían ser una herramienta eficaz y útil para implementar nuevos protocolos de *screening* de SAHS. Complementariamente, los métodos de reconocimiento automático de patrones propuestos permiten maximizar la capacidad diagnóstica de los registros de SpO<sub>2</sub> portátiles no supervisados.

## Agradecimientos

Este estudio ha sido financiado por el proyecto 265/2012 de la Sociedad Española de Neumología y Cirugía Torácica (SEPAR) y por el proyecto VA059U13 de la Consejería de Educación de la Junta de Castilla y León.

## Referencias

- [1] Flemons WW, Littner MR, Rowley JA, Gay P, Mc Andersonc W, Hudgel DW, McEvoy D, Loubé D. Home diagnosis of sleep apnea: a systematic review of the literature. *Chest*, vol 124, 2003, pp 1543-79
- [2] Fernández I, Heili S, Peces-Barba G. Síndrome de apnea obstructiva del sueño. *Patología Respiratoria. Manual de actuación*. Ed. Ergon, 2004
- [3] Whitelaw WA, Brant RF, Flemons WW. Clinical usefulness of home oximetry compared with polysomnography for assessment of sleep apnea. *Am. J. Respir. Crit. Care Med.*, vol 171, 2005, pp 188-93
- [4] Iber C, Ancoli-Israel S, Chesson AL, Quan SF, for the American Academy of Sleep Medicine. The AASM Manual for the scoring of sleep and associated events. Rules, terminology and technical specifications, 1ST Ed., American Academy of Sleep Medicine, 2007



**Figura 4.** Distribución de valores a la salida de la red neuronal MLP de regresión en la población de test

- [5] Alvarez D, Hornero R, Marcos JV, Wessel N, Penzel T, Glos M, Del Campo F. Assessment of feature selection and classification approaches to enhance information from overnight oximetry in the context of sleep apnea diagnosis. *Int. J. Neural Syst.*, vol 23, 2013, pp 1-18
- [6] Gutiérrez-Tobal GC, Álvarez D, Marcos JV, Del Campo F, Hornero R. Pattern recognition in airflow recordings to assist in the sleep apnoea-hypopnoea syndrome diagnosis. *Medical & Biological Engineering & Computing*, vol 51, 2013, pp 1367-80
- [7] Bishop CM. *Neural networks for pattern recognition*. Oxford University Press, 1995
- [8] Álvarez D, Hornero R, Marcos JV, Del Campo F. Multivariate analysis of blood oxygen saturation recordings in obstructive sleep apnea diagnosis. *IEEE Trans. Biomed. Eng.*, vol 57, 2010, pp 2816-24
- [9] Zamarrón C, Gude F, Barcala J, Rodríguez JR, Romero PV. Utility of oxygen saturation and heart rate spectral analysis obtained from pulse oximetric recordings in the diagnosis of sleep apnea syndrome. *Chest*, vol 123, 2003, pp 1567-76
- [10] Hosmer DW, Lemeshow S. *Applied Logistic Regression*. John Wiley & Sons, 1989
- [11] Guyon I, Elisseeff A. An introduction to variable and feature selection. *Journal of Machine Learning Research*, vol 3, 2003, pp 1157-82
- [12] Jobson JD. *Applied multivariate data analysis. Volume II: Categorical and multivariate methods*. Springer-Verlag, 1991
- [13] Gurubhagavatula I, Fields BG, Morales CR, Hurley S, Pien GW, et al. Screening for severe obstructive sleep apnea syndrome in hypertensive outpatients. *Journal of Clinical Hypertension*, vol 25, 2013, pp 279-88
- [14] Torres G, Sánchez de la Torre M, Martínez-Alonso M, Gómez S, Sacristán O, Cabau J, Berbé F. Use of ambulatory home pressure monitoring for the screening of obstructive sleep apnea. *Journal of Clinical Hypertension*, 2015, *In press*
- [15] Crowley KE, Rajaratnam SMW, Shea SA, Epstein LJ, Czeisler CA, Lockley SW. Evaluation of a single-channel nasal pressure device to assess obstructive sleep apnea risk in laboratory and home environments. *J Clin Sleep Med*, vol 9, 2013, pp 109-16

