

# Procesado Automático de Señales Cardiorrespiratorias para la Caracterización de Pacientes con COVID-19 Persistente

M. Fernández Vaquerizo<sup>1</sup>, F. Moreno Torrero<sup>2</sup>, G. C. Gutiérrez Tobal<sup>3,4</sup>, T. Ruiz Albi<sup>2,3,4</sup>, R. Hornero Sánchez<sup>3,4</sup>, F. del Campo Matía<sup>3,4</sup>, D. Álvarez González<sup>3,4</sup>

<sup>1</sup> Departamento de Innovación e Investigación, Fundació Privada Hospital Asil de Granollers, Granollers, Barcelona, España, mariafernandezvaquerizo@gmail.com

<sup>2</sup> Servicio de Neumología, Hospital Universitario Río Hortega, Valladolid, España, {fmorennot, truzal}@saludcastillayleon.es

<sup>3</sup> Grupo de Ingeniería Biomédica, Universidad de Valladolid, Valladolid, España, {gonzalocesar.gutierrez, roberto.hornero, felix.campo, daniel.alvarez.gonzalez}@uva.es

<sup>4</sup> CIBER de Bioingeniería, Biomateriales y Nanomedicina, Instituto de Salud Carlos III, Valladolid, España

## Resumen

*El síndrome post-COVID o COVID persistente (SPC-CP) se define como la continuidad de los síntomas más allá de las 12 semanas posteriores a una infección por COVID-19, siendo actualmente uno de los mayores problemas relacionados con la enfermedad tras el fin de la pandemia. El objetivo del estudio fue doble: (i) identificar los índices autonómicos, oximétricos y de acoplamiento cardiorrespiratorio con mayor asociación con el SPC-CP y (ii) diferenciar a pacientes con afectación autonómica potencialmente asociada a síntomas persistentes de otros sin SPC-CP. Se estudiaron 3 grupos de pacientes (control no-COVID, COVID, SPC-CP), procesando individual y conjuntamente la variabilidad del ritmo cardiaco (HRV), oximetría y respiración derivada del ECG, empleando enfoques complementarios (tiempo vs. frecuencia y lineal vs. no lineal). Se implementó una etapa de selección de variables basada en algoritmos genéticos y se diseñó un clasificador multiclase mediante una red neuronal perceptrón multicapa (MLP). Se formó una población de estudio compuesta por 83 sujetos (25 control, 25 COVID, 33 SPC-CP), siendo la presencia de comorbilidad respiratoria, potencia a altas frecuencias en la señal HRV, magnitud de la coherencia espectral del acoplamiento y puntuación del COPD Assessment Test (CAT) de calidad de vida las variables con mayor capacidad discriminante. La red MLP alcanzó 75% Acc, 0.621 kappa y 76.52% macro F1-score en la identificación de las 3 clases de pacientes, así como 90% Acc (87.50% Se y 91.67% Sp) en la discriminación de SPC-CP. Estos hallazgos sugieren que el análisis automático de la actividad cardiorrespiratoria podría aportar información útil en la caracterización y manejo de pacientes con SPC-CP.*

## 1. Introducción

El virus SARS-CoV-2 es el agente causal del COVID-19, síndrome respiratorio descubierto en China en diciembre de 2019 [1], siendo declarada pandemia global en marzo de 2020. Tras la finalización de la declaración de emergencia sanitaria internacional, una de las principales preocupaciones en torno al COVID-19 es la persistencia de los síntomas, que podrían tener un efecto perjudicial a largo plazo. Entre el 10-35% de los pacientes que se han recuperado refieren la persistencia de al menos un síntoma, siendo los más frecuentes la fatiga y la disnea [2]. El *National Institute for Health and Care Excellence* define como síndrome post-COVID o COVID persistente

(SPC-CP) a los síntomas que continúan o se desarrollan más allá de las 12 semanas posteriores a la infección por COVID-19 y que no se pueden explicar mediante un diagnóstico alternativo, exhibiendo estos pacientes deterioro de la estructura y función de múltiples órganos [1, 3]. La prevalencia y la sintomatología del SPC-CP son muy heterogéneas, destacando la fatiga (52% de los casos) y otros síntomas cardiorrespiratorios (30-42% de los casos) y neurológicos (40% de los casos) [2, 4].

Diferentes estudios han demostrado que el SARS-CoV-2 afecta al sistema autónomo [3, 4]. Se ha observado que los pacientes con SPC-CP exhiben una desregulación de la variabilidad del ritmo cardiaco (HRV) que podría explicar algunos síntomas [4]. Sin embargo, las comparaciones se han realizado con respecto a controles puros sin antecedentes de COVID-19, sin incluir a pacientes que hayan tenido enfermedad no persistente.

El presente trabajo se desarrolló bajo la hipótesis de que el análisis de la actividad cardiaca, respiratoria y su interacción (acoplamiento cardiorrespiratorio, CRC) podría aportar información útil en el manejo de pacientes con SPC-CP. El objetivo general del estudio fue doble: (i) caracterizar los patrones cardiorrespiratorios de pacientes con y sin diagnóstico de SPC-CP para identificar los índices autonómicos, oximétricos y de acoplamiento con mayor asociación con la persistencia de síntomas y (ii) diferenciar a pacientes con alteración cardiorrespiratoria potencialmente asociada al SPC-CP de otros sujetos sin diagnóstico de SPC-CP. Para ello, se propone analizar las señales de HRV, oximetría y CRC mediante técnicas de procesado automático y *machine learning*.

## 2. Diseño del estudio. Sujetos y señales

Estudio prospectivo, longitudinal y observacional. Población de estudio formada por pacientes con diagnóstico previo de COVID-19 que acudieron a las consultas de seguimiento del Servicio de Neumología del Hospital Universitario Río Hortega de Valladolid, entre febrero y abril de 2023. Los criterios de inclusión fueron: diagnóstico previo de COVID-19, ambos sexos y edad comprendida entre 18 y 80 años. Se consideró SPC-CP a los pacientes con diagnóstico previo de COVID-19 y

persistencia de síntomas superior a 12 semanas [3]. Paralelamente, se incorporaron al estudio sujetos de control reclutados entre los pacientes que acudieron a las consultas generales de Neumología en el mismo periodo de tiempo, creando los siguientes grupos de estudio:

- **Grupo 0:** pacientes sin diagnóstico previo de infección por COVID-19 (grupo de control puro).
- **Grupo 1:** pacientes con diagnóstico previo confirmado de COVID-19 que actualmente no presentan sintomatología (grupo de control COVID).
- **Grupo 2:** pacientes con diagnóstico clínico confirmado de SPC-CP (grupo de estudio SPC-CP).

Se definieron los siguientes criterios de exclusión: presentar patologías cardíacas o respiratorias graves con prescripción de medicación que afectase al sistema nervioso autónomo. El estudio fue aprobado por el Comité de Ética del Hospital Universitario Río Hortega (CEIm 23-PI010) y se solicitó la firma de un consentimiento informado a todos los participantes.

Para cada paciente se recopilaron variables sociodemográficas (edad, sexo), antropométricas (índice de masa corporal, tensión arterial), antecedentes (hábito tabáquico, comorbilidades, vacunación, hospitalización), clínicas (presencia y duración de síntomas) y de calidad de vida (disnea: escala de disnea modificada del Medical Research Council (mMRC); afectación de las actividades de la vida diaria: test CAT). Igualmente, se registraron simultáneamente las señales de electrocardiograma (ECG) y saturación de oxígeno en sangre ( $SpO_2$ ) durante 7 minutos, con el paciente sentado y en reposo.

El ECG se registró a una frecuencia de muestreo (fs) de 1 KHz (Biosignalsplux, Plux Wireless Biosignals S.A., Lisboa, Portugal), obteniendo de él la señal HRV (serie temporal del intervalo RR) y la señal de respiración derivada del ECG (envolvente de los complejos QRS). La señal de  $SpO_2$  se adquirió a una fs de 1 Hz (NONIN WristOx2 3150, Nonin Medical Inc., MN, EE.UU.).

### 3. Metodología

En la Figura 1 se muestran las particiones de la población de estudio utilizadas para implementar las etapas de diseño, optimización y validación del sistema de ayuda a la toma de decisiones. El 75% de la población (*hold-out* 1) se empleó para seleccionar las características óptimas y entrenar (*leave-one-out*) la red neuronal. El 25% restante se reservó para realizar un test independiente del modelo final. La segunda partición (*hold-out* 2 en 75% anterior) se empleó para optimizar (*k-fold*) los hiperparámetros de las múltiples redes MLP evaluadas durante las llamadas a la función de ajuste (maximización de Acc) en el proceso de selección de variables basado en algoritmos genéticos.

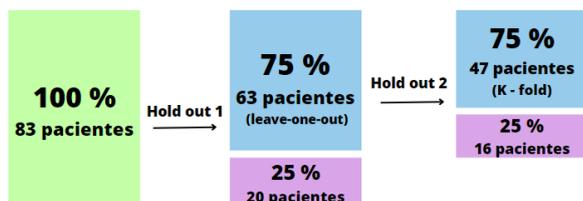


Figura 1. Partición de la población bajo estudio.

#### 3.1. Preprocesado de las señales biomédicas

Para cada una de las señales biomédicas bajo estudio se realizó un preprocesado específico, descartando los pacientes con tiempo de registro válido inferior a 5 minutos. Para la señal HRV, se descartaron las muestras con valores  $<0.33$  segundos (falso latido),  $>1.50$  segundos (latido perdido) o con diferencias  $>0.66$  entre intervalos RR consecutivos, interpolando mediante *splines* cúbicos para obtener una frecuencia de muestreo uniforme de 3.41 Hz [5]. Para la señal de  $SpO_2$ , se eliminaron las muestras perdidas asociadas a movimientos del paciente, concatenando los segmentos válidos.

#### 3.2. Extracción de características

##### 3.2.1. Análisis individual de las señales HRV y $SpO_2$

Se estimó la densidad espectral de potencia (PSD) mediante el método de Welch para las señales de HRV y  $SpO_2$ . De cada espectro se calculó la potencia total (PT), la frecuencia mediana (MF) y la entropía espectral (SE). Se realizó un análisis no lineal para obtener información complementaria sobre la dinámica de ambas señales [6]: entropía muestral (SampEn), medida de la tendencia central (CTM) y complejidad de Lempel-Ziv (LZC).

Igualmente, de la señal HRV se extrajeron los índices tradicionales de modulación autonómica: intervalo NN promedio (AVNN), desviación estándar de los intervalos NN (SDNN), valor cuadrático medio de las diferencias entre intervalos NN consecutivos (RMSSD), potencia relativa a muy bajas ( $P_{VLF}$ ), bajas ( $P_{LF}$ ), y altas frecuencias ( $P_{HF}$ ) y balance simpático-vagal ( $P_{LF/HF}$ ). Para la señal de  $SpO_2$ , se calcularon los índices oximétricos básicos: valor medio, basal y mínimo; índices de eventos de desaturación del 2, 3 y 4%; tiempo acumulado con valores por debajo de 90% y 95%.

##### 3.2.2. Análisis conjunto: acoplamiento cardiorrespiratorio

Para caracterizar el CRC se emplearon diferentes enfoques complementarios: dominio temporal, secuencia de correlación (MCC); frecuencia, magnitud de la coherencia espectral (MSC); no lineal, información mutua (MI). Además, se calcularon dos índices empleados comúnmente para cuantificar la magnitud de la modulación del ECG por la respiración (RSA): promedio de la diferencia entre el intervalo RR más largo y el más corto para cada respiración; y promedio de las diferencias entre la frecuencia cardíaca y la frecuencia respiratoria para cada ciclo respiratorio.

#### 3.3. Selección de características

La selección del subconjunto óptimo de variables se realizó mediante algoritmos genéticos (GAs). El conjunto de partida (espacio de búsqueda) estuvo compuesto por 58 variables. La función de ajuste a optimizar fue la correspondiente al rendimiento predictivo (Acc mediante *k-fold* en el 75% de pacientes del *hold-out* 2) de una red MLP multiclasificada diseñada para identificar los 3 grupos de pacientes.

### 3.4. Clasificación de características

Se utilizó una red MLP con una capa oculta y la función de activación *softmax*. Los pesos fueron optimizados mediante el algoritmo *scaled conjugate gradient* con un número máximo de 100 iteraciones. El modelo de red final fue entrenado en el conjunto de entrenamiento completo (*hold-out 1*) empleando los valores óptimos del número de nodos (*Nh*) y tasa de aprendizaje ( $\alpha$ ) obtenidos mediante una estrategia *leave-one-out cross-validation*.

### 3.5. Análisis estadístico

Se empleó el test de Kruskal-Wallis para la búsqueda de diferencias estadísticas entre los tres grupos de estudio. Se aplicó un análisis *post-hoc* mediante el test de Mann-Whitney para determinar entre qué grupos 2-a-2 se observaban diferencias significativas. Para las variables categóricas se empleó el test Chi<sup>2</sup>. Se aplicó la corrección de Fisher, con un umbral de significación de 0.05.

La capacidad predictiva del clasificador multiclase se evaluó en términos de precisión global (Acc), índice Kappa y *macro-F1 score*. Para valorar el enfoque binario se emplearon las métricas clásicas de sensibilidad (Se), especificidad (Sp), valor predictivo positivo (PPV) y negativo (NPV), razón de verosimilitud positiva (LR+) y negativa (LR-) y precisión (Acc).

## 4. Resultados

### 4.1. Población bajo estudio

Un total de 105 pacientes cumplieron los criterios de inclusión y firmaron el consentimiento. Se utilizó la distancia robusta de Mahalanobis con una sensibilidad del 5% para la detección de *outliers*. La Figura 2 muestra el diagrama de flujo de los participantes. En la *Tabla 1* se muestran las variables socio-demográficas básicas de la población bajo estudio.

El análisis descriptivo univariante arrojó diferencias significativas ( $p < 0.05$ ) para la tensión arterial diastólica (control vs. COVID), disnea y calidad de vida (control vs. SPC-CP), CTM de la señal SpO<sub>2</sub> (control vs. COVID y COVID vs. SPC-CP) y el máximo de la MCC (COVID vs. SPC-CP).

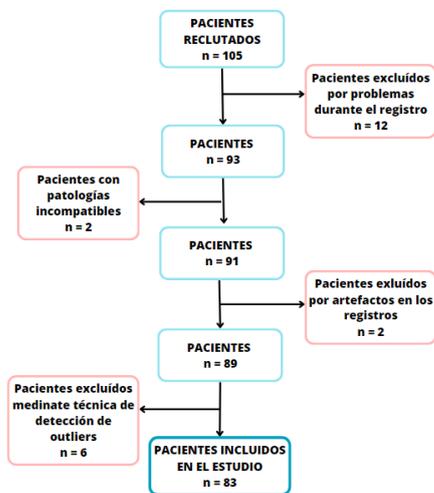


Figura 2. Diagrama de flujo de la población bajo estudio.

Variables	Sujetos	Control	COVID	SPC-CP	p-valor
Sujetos	83 (100%)	25 (30%)	25 (30%)	33 (40%)	-
Sexo (masculino)	48 (58%)	13 (16%)	17 (20%)	18 (22%)	0.4596
Edad (años)	54.00 [16.00]	53 [17.75]	52.00 [16.25]	56.00 [14.50]	0.2363
IMC	28.03 [6.81]	26.89 [6.32]	30.90 [5.50]	27.25 [5.21]	0.1564

IMC: índice de masa corporal. Datos presentados como número de sujetos (%) o mediana [rango intercuartil].

Tabla 1. Población bajo estudio.

### 4.2. Selección de características

Independientemente de su capacidad individual, las variables seleccionadas mediante GAs por su mayor capacidad discriminante conjunta fueron (4/58, 6.9%): presencia de comorbilidad respiratoria (episodios apneicos), P<sub>HF</sub>, MSC entre las señales HRV y respiración derivada, y test CAT de calidad de vida.

### 4.3. Red neuronal perceptrón multicapa

Los hiperparámetros óptimos de la red MLP final fueron  $Nh = 5$  en la capa oculta y  $\alpha = 1$ . En la Figura 3 se muestra la matriz de confusión en el conjunto independiente de test. Se observa que los grupos con mayor discriminación fueron los de COVID y SPC-CP. La red MLP multiclase alcanzó una precisión global del 75.00%, índice kappa de 0.621 y *macro-F1 score* del 76.52% en el conjunto de test.

La Figura 3 muestra la matriz de confusión obtenida al discriminar a los pacientes con SPC-CP de aquellos sin síntomas persistentes (No COVID + COVID). La Tabla 2 resume las métricas de rendimiento predictivo derivadas, destacando una Acc del 90.0% con un par Se-Sp muy balanceado (87.50 vs. 91.67) y LR+ elevado (10.50).

**Matriz de confusión multiclase**

True Class	CONTROL	3	2	1
	COVID	1	5	
	SPC-CP	1		7
		CONTROL	COVID	SPC-CP

Predicted Class

Figura 3. Matriz de confusión multiclase.

**Matriz de confusión binaria**

True Class	NO SPC-CP	11	1
	SPC-CP	1	7
		NO SPC-CP	SPC-CP

Predicted Class

Figura 4. Matriz de confusión de clasificación binaria.

Se	Sp	PPV	NPV	LR+	LR-	Acc
87.50 %	91.67 %	87.50 %	91.67 %	10.50	0.14	90.00 %

Tabla 2. Métricas de rendimiento al discriminar SPC-CP.

## 5. Discusión

En este estudio se ha analizado exhaustivamente la actividad cardiorrespiratoria y las características clínicas de pacientes con SPC-CP, evaluando sus diferencias con grupos de pacientes COVID sin persistencia de síntomas y sujetos sin antecedentes de la enfermedad. La principal novedad del estudio radica en el análisis del acoplamiento cardiorrespiratorio desde enfoques complementarios, así como en el diseño de una red neuronal para identificar pacientes con SPC-CP, que podrían llegar a las consultas sin un diagnóstico definido. Hasta donde conocemos, no existen estudios donde se analice la influencia del SPC-CP sobre la actividad CRC, ni se han incluido grupos de control con pacientes COVID sin síntomas persistentes.

Respecto a las variables clínicas, en el caso de la tensión arterial diastólica (TAD), varios estudios reportan un incremento de la presión sanguínea [4, 7, 8], lo que concuerda con los resultados obtenidos. Respecto a los dos cuestionarios clínicos evaluados, la escala de disnea mMRC y el cuestionario de calidad de vida CAT, ambos presentaron diferencias entre los dos grupos control y el grupo SPC-CP.

Con respecto a la señal de SpO<sub>2</sub>, se decidió analizarla debido a que la infección por COVID-19 se caracteriza por una respuesta inmunoinflamatoria en los tejidos pulmonares [7, 8]. Las diferencias significativas observadas en el CTM de esta señal, con mayor variabilidad en el grupo SPC-CP, confirma su elección. Aunque no se encontraron diferencias significativas al analizar las características derivadas de la señal HRV, sí que se seleccionó la potencia a altas frecuencias (relacionada con la respiración) como parte del subconjunto óptimo de variables, lo que demuestra su relevancia y complementariedad. De forma similar, se encontraron diferencias en el acoplamiento cardiorrespiratorio (MCC), donde el grupo SPC-CP mostró una magnitud menor (menor sincronización entre la actividad respiratoria y la cardíaca).

Con respecto a las variables óptimas incluidas en el modelo final (AOS, P<sub>HF</sub>, MSC y CAT), varios estudios apuntan a que el SPC-CP está relacionado con alteraciones del sueño [9, 10]. Igualmente, la variable P<sub>HF</sub> adquiere sentido al estar relacionada con el control parasimpático de la frecuencia cardíaca, que a su vez refleja las variaciones causadas por la respiración [11]. Por otro lado, la MSC ha demostrado ser útil para evaluar la sincronización entre diferentes regiones del sistema cardiovascular y el respiratorio [12]. Con respecto al test CAT, es conocida la relación entre los síntomas y la afectación de la calidad de vida, lo que concuerda con los valores más elevados (mayor afectación) en SPC-CP.

Se observó que los grupos que mejor discriminaba la red MLP eran los de COVID y SPC-CP. Los resultados sugieren que existen características clínicas particulares de los pacientes SPC-CP que se maximizan al diferenciarlos de los pacientes agrupados sin síntomas independientemente de que hayan pasado el COVID o no. En la literatura no hay estudios que hayan propuesto

modelos automáticos similares para ayudar en el manejo de los pacientes con SPC-CP.

Es necesario indicar algunas limitaciones que podrían condicionar los resultados. El grupo de pacientes de control podría tener otras patologías, incluso haber padecido COVID-19 asintomático. Además, el tamaño de la muestra podría haber sido mayor.

## 6. Conclusiones

Los resultados obtenidos demuestran que existen diferencias en la oximetría (CTM), acoplamiento cardiorrespiratorio (MCC), TAD y en la afectación a la calidad de vida (test de disnea y CAT) entre pacientes con y sin SPC-CP, lo que sugiere la influencia de esta condición en los sistemas cardíaco y respiratorio. Igualmente, existe información complementaria en las variables AOS, CAT, P<sub>HF</sub> y MSC que permite identificar con elevada precisión a pacientes con SPC-CP, lo que sugiere que los métodos de procesamiento automático de señal y aprendizaje computacional podrían ser útiles en el manejo de estos pacientes en la práctica clínica.

## 7. Agradecimientos

Este estudio ha sido financiado parcialmente por el CIBER-BBN (CB19/01/00012) del Instituto de Salud Carlos III y por el Proyecto CPP2022-009735 financiado por MICIU/AEI/10.13039/501100011033 y por la Unión Europea NextGenerationEU/ PRTR. D. Álvarez es beneficiario de una ayuda Ramón y Cajal (RYC2019-028566-I) del Ministerio de Ciencia e Innovación - AEI, cofinanciada por el FSE.

## Referencias

- [1] Zhu N, Zhang D, Wang W, et al. A Novel Coronavirus from Patients with Pneumonia in China, 2019. *N Engl J Med* 2020;382:727–33.
- [2] Pavli A, Theodoridou M, Maltezou HC. Post-COVID Syndrome: Incidence, Clinical Spectrum, and Challenges for Primary Healthcare Professionals. *Arch Med Res* 2021;52:575–81.
- [3] Crook H, Raza S, Nowell J, Young M, Edison P. Long covid-mechanisms, risk factors, and management. *BMJ* 2021;374:n1648.
- [4] Salem AM, Yar T, Al Eid M, et al. Post-Acute Effect of SARS-CoV-2 Infection on the Cardiac Autonomic Function. *Int J Gen Med* 2022;15:7593–603.
- [5] Penzel T, Kantelhardt JW, Grote L, Peter J-H, Bunde A. Comparison of detrended fluctuation analysis and spectral analysis for heart rate variability in sleep and sleep apnea. *IEEE Trans Biomed Eng* 2003;50:1143–51.
- [6] Alvarez D, Hornero R, Marcos JV, et al. Assessment of feature selection and classification approaches to enhance information from overnight oximetry in the context of apnea diagnosis. *Int J Neural Syst* 2013;23:1350020.
- [7] Al-Hadrawi DS, Al-Rubaye HT, Almulla AF, Al-Hakeim HK, Maes M. Lowered oxygen saturation and increased body temperature in acute COVID-19 largely predict chronic fatigue syndrome and affective symptoms due to Long COVID: A precision nomothetic approach. *Acta Neuropsychiatr* 2023;35:76–87.
- [8] Al-Hakeim HK, Al-Rubaye HT, Al-Hadrawi DS, Almulla AF, Maes M. Long-COVID post-viral chronic fatigue and affective symptoms are associated with oxidative damage, lowered antioxidant defenses and inflammation: a proof of concept and mechanism study. *Mol Psychiatry* 2023;28:564–78.
- [9] Scarpelli S, De Santis A, Alfonsi V, et al. The role of sleep and dreams in long-COVID. *J Sleep Res* 2023;32:e13789.
- [10] Alzueta E, Perrin PB, Yuksel D, et al. An international study of post-COVID sleep health. *Sleep Health* 2022;8:684–90.
- [11] Stein PK, Pu Y. Heart rate variability, sleep and sleep disorders. *Sleep Med Rev* 2012;16:47–66.
- [12] Nolte G, Bai O, Wheaton L, Mari Z, Vorbach S, Hallett M. Identifying true brain interaction from EEG data using the imaginary part of coherency. *Clin Neurophysiol* 2004;115:2292–307.