

# Relación entre la Actividad Física y el Ritmo Cardíaco Registrados por un Wearable de Pulsera en Pacientes con Depresión

A. Barquero<sup>1,\*</sup>, E. García<sup>2,3</sup>, S. Kontaxis<sup>1</sup>, S. Siddi<sup>4</sup>, N. Cummins<sup>5</sup>, S. Vairavan<sup>6</sup>, F. Matcham<sup>5,7</sup>, M. Hotopf<sup>5,8</sup>, F. Lamers<sup>9,10</sup>, B. Penninx<sup>9,10</sup>, R. Dobson<sup>5</sup>, V. Narayan<sup>6</sup>, R. Bailón<sup>1,2</sup>, P. Armañac-Julián<sup>1,2</sup>, the RADAR-CNS consortium

<sup>1</sup>Grupo BSICoS, I3A, IIS Aragón, Universidad de Zaragoza, España, 761145@unizar.es; <sup>2</sup>Centro de Investigación Biomédica en Red – BBN (CIBER-BBN), España; <sup>3</sup>Microelectrónica y Sistemas Electronicos, Universidad Autónoma de Barcelona, C; <sup>4</sup>Parc Sanitari Sant Joan de Déu, Fundació Sant Joan de Déu, CIBERSAM, Universitat de Barcelona, España; <sup>5</sup>King's College London, Institute of Psychiatry, Psychology and Neuroscience, London, UK; <sup>6</sup>Research and Development Information Technology, Janssen Research & Development, LLC, Titusville, NJ, USA; <sup>7</sup>School of Psychology, University of Sussex, Falmer, UK; <sup>8</sup>South London and Maudsley NHS Foundation Trust, UK; <sup>9</sup>Department of Psychiatry, Amsterdam UMC, Vrije Universiteit, the Netherlands; <sup>10</sup>Amsterdam Public Health Research Institute, the Netherlands.

## Resumen

*Este estudio investiga una nueva forma de monitoreo en pacientes con depresión mediante el uso de un smartwatch Fitbit, dentro del proyecto europeo RADAR-CNS. Durante dos años, se recopilaron datos de frecuencia cardíaca (HR) y actividad física para identificar patrones en pacientes con distintos niveles de depresión. Las hipótesis eran: (1) los pacientes con depresión o recaídas realizan menos actividad física que aquellos en recuperación; (2) los pacientes con depresión severa tienen una recuperación cardíaca más lenta, lo que sugiere un impacto cardiovascular. Para evaluar la recuperación tras la actividad física, se utilizaron varios indicadores: HR máximo y mínimo, capacidad cardíaca (diferencia entre ambos), T95 (tiempo para alcanzar el 95% de la recuperación) y número de ventanas de actividad (NW). Estos biomarcadores se correlacionaron con los síntomas clínicos de depresión. Los resultados en 217 pacientes indicaron que, tras ejercicios intensos, aquellos con depresión severa realizaron menos actividad física (NW: correlación de 0.0182 a 0.9663 con la severidad de la depresión) y tuvieron una recuperación cardíaca más lenta (HR: 0.0144 a 0.9572 bpm). El T95 mostró tiempos de recuperación prolongados en los pacientes con depresión severa, respaldando un posible impacto cardiovascular. Sin embargo, estas diferencias no fueron estadísticamente significativas, sugiriendo que el nivel de depresión no afecta de manera concluyente la recuperación cardíaca. Finalmente, se concluye que los datos fisiológicos deben complementarse con evaluaciones clínicas para apoyar el manejo de pacientes con depresión.*

## 1. Introducción

El Trastorno Depresivo Mayor (TDM) es una enfermedad grave que afecta al 6% de la población adulta mundial, asociándose con consecuencias como discapacidad, disminución de la calidad de vida, mortalidad prematura y suicidio, siendo más prevalente en mujeres que en hombres [1]. Su etiología es multifactorial, y dada su compleja sintomatología, que incluye alteraciones en el estado de ánimo y el sueño, muchas de ellas relacionadas con el sistema nervioso autónomo, la monitorización mediante dispositivos portátiles (wearables) ha surgido como una herramienta prometedora. Los wearables permiten el registro continuo de parámetros fisiológicos, como la

frecuencia cardíaca (HR), y la actividad física, lo que podría ayudar a identificar cambios tempranos en el estado de salud que podrían pasar desapercibidos en evaluaciones clínicas esporádicas, así como la realización de test con respuestas subjetivas, basadas en la percepción del individuo (Patient Health Questionnaire, PHQ-8) pueden ayudar a los equipos terapéuticos a conocer síntomas no detectables en un entorno extrahospitalario, y desarrollar un tratamiento adaptado a las necesidades de cada paciente [2]. Esta tecnología facilita una monitorización objetiva y en tiempo real, ofreciendo la posibilidad de predecir recaídas y optimizar el tratamiento, mejorando la calidad de vida de los pacientes.

Este estudio combina HR y actividad física para analizar la recuperación de la HR (HRR) tras la actividad física en pacientes con depresión. La hipótesis básica de este estudio es que la depresión se ha relacionado con alteraciones del sistema nervioso autónomo [3]. Estas alteraciones están asociadas con mayor riesgo cardiovascular, y también se ha visto que los pacientes con depresión tienen un mayor riesgo cardiovascular que la población normal [4]. Por ende, el objetivo principal del trabajo es evaluar si la HRR al ejercicio durante la vida diaria será más lenta en aquellos con depresión severa, indicando afectación cardiovascular y relacionarla con indicadores de depresión, así como la gravedad de esta (depresión severa o no severa).

## 2. Materiales y Métodos

### 2.1. Base de Datos

Se han utilizado los datos de las series de HR y Pasos de 217 pacientes con TDM, pertenecientes al proyecto RADAR-CNS [5]. Los participantes han utilizado tecnologías de medición remotas y, a la par, se ha realizado un seguimiento de su evolución clínica. Para ello, todos han utilizado una pulsera Fitbit durante su vida cotidiana, que ha recogido datos de forma continua de su HR, como el resultado de los múltiples test de PHQ-8 cada 2 semanas entre otros parámetros durante un mínimo de 11 meses a

un máximo de 24 meses. La escala de depresión PHQ-8 es una herramienta validada para diagnosticar y medir la gravedad de trastornos depresivos en estudios clínicos, se realiza cada 2 semanas, y se utilizará como referencia para evaluar la severidad de la depresión. Está compuesta por una serie de preguntas evaluadas en una escala de '0' (nunca) a '3' (casi todos los días), la puntuación total se obtiene sumando las respuestas. Una puntuación de 10 o más indica depresión mayor, mientras que 20 o más sugiere depresión mayor severa [2].

### 1.2. Respuesta Promedio del HR

En este estudio, se aplica la metodología de la prueba de marcha de 6 minutos (6MWT) para detectar pruebas de caminata no intencionales en los datos de pasos de los pacientes [6]. Se utilizan ventanas deslizantes de 6 minutos con 1 minuto de superposición, dado que las actividades diarias presentan variaciones en la cadencia de pasos. Para la monitorización precisa de la actividad física con dispositivos portátiles, es esencial identificar los momentos de inicio (onset) y finalización (offset) de las ventanas de actividad (véase Figura 1). El onset marca el comienzo de la actividad física cuando la cadencia de pasos supera un umbral predefinido, mientras que el offset indica su final, cuando la cadencia disminuye por debajo del nivel establecido. Estos puntos se determinan a partir de los datos de pasos registrados por el dispositivo Fitbit, lo que permite evaluar cómo la actividad afecta la HR.

En el análisis, se establecen tres niveles para detectar ventanas de actividad continua de mínimo 6 minutos: (1) más de 0 pasos por minuto para definir movimiento continuo; (2) al menos 60 pasos por minuto para identificar actividad moderada estandarizada [6]; y (3) un umbral personalizado basado en la mediana de pasos de cada paciente, detectando así actividad intensa individualizada. Además, el incremento en la HR depende del esfuerzo metabólico, que varía entre individuos. Para extraer patrones en los datos de pasos y HR, se utiliza el método Phase-rectified signal averaging (PRSA), diseñado para detectar oscilaciones cuasiperiódicas en señales ruidosas. En este caso, se emplea una versión bivariante del PRSA para estudiar la interacción entre los pasos y la HR [7]. A través de los offsets definidos en las ventanas de actividad de la serie de pasos, se observa cómo se comporta la HRR (véase Figura 2). Las series de pasos y HR fueron remuestreadas a intervalos de 5 segundos para garantizar la coherencia en el análisis. Finalmente, la curva BPRSA se obtiene promediando todas las ventanas de actividad segmentadas en la HR durante un intervalo determinado, como por ejemplo dos semanas, y clasificando la respuesta cardíaca según la intensidad de la actividad física detectada.

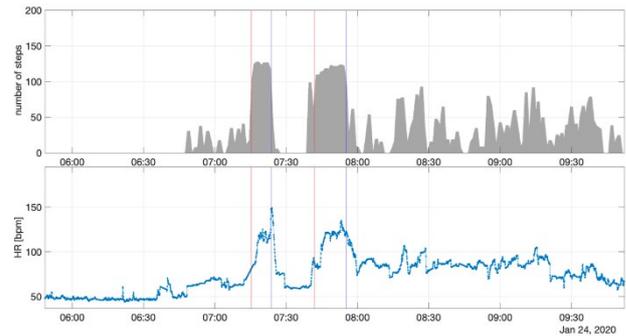


Figura 1. Onset y offset definidos sobre la serie de pasos, y llevados a la serie de HR sincronizada. En rojo se muestra los onset, y en azul los offset.

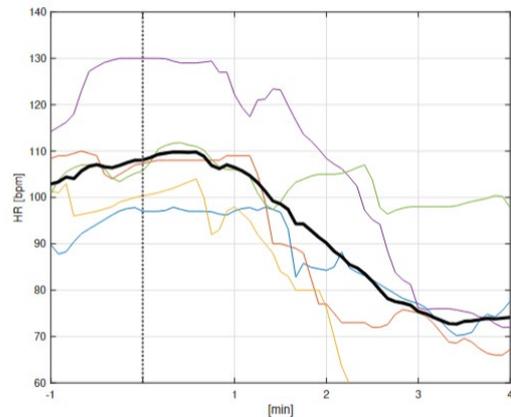


Figura 2. Curva de respuesta de recuperación promedio del HR, ante el offset de actividad física.

Se definen por tanto las siguientes variables para caracterizar la HRR: Número de ventanas (NW), HR máxima en la curva promedio ( $HR_{\max}$ ), mínimo de HR en la curva promedio ( $HR_{\min}$ ), Capacidad cardíaca calculada como el máximo menos el mínimo de la HR en la curva promedio, y tiempo en llegar al 95% del  $HR_{\min}$  en la curva promedio ( $T_{95}$ ) (ver Figura 3). También se han definido otras variables como la duración de las ventanas de actividad (DVA), el HR a cada minuto después del offset, y la capacidad a cada minuto (por ejemplo,  $HRR_{1\min}$  calculada como el  $HR_{\max} - HR_{1\min}$ ).

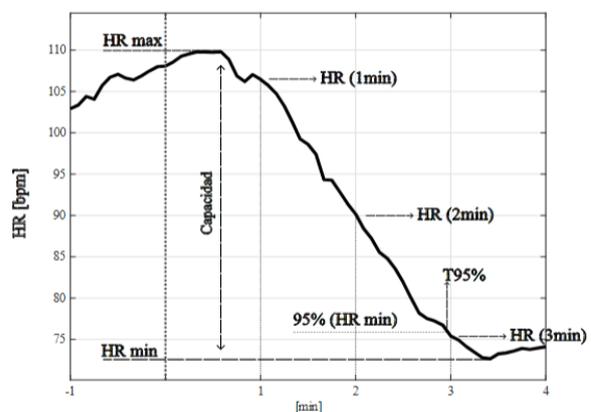


Figura 3. Variables definidas en la curva promedio.

### 1.3. Análisis estadístico

Para evaluar si existen diferencias significativas entre los niveles de depresión para cada una de las variables, se utiliza el test U de Mann-Whitney. Como cada paciente tiene varias mediciones del PHQ-8 a lo largo del estudio, inicialmente habría que seleccionar una única muestra aleatoria para realizar el test de Wilcoxon. Sin embargo, esto podría llevar a perder información valiosa, por lo que se emplea la técnica de Bootstrap para aprovechar todos los datos disponibles y obtener resultados más robustos.

El método Bootstrap genera múltiples muestras aleatorias con reemplazo, permitiendo capturar la variabilidad de los datos [8]. Este enfoque proporciona estimaciones más precisas al realizar el test U de Mann-Whitney. En este caso, se realizaron 2000 réplicas mediante Bootstrap, lo que permitió calcular el p-valor de cada variable relacionada con los dos grupos de depresión. Finalmente, se calcularon los intervalos de confianza del 5 % y el 95 % para los p-valores, lo que aporta una estimación más precisa del rango en el que se esperan los valores de p para cada variable.

Para explorar las relaciones no lineales entre las variables definidas y el nivel de depresión medido por el PHQ-8, se realiza un análisis de correlación de Spearman. Este método es adecuado para capturar asociaciones entre variables, incluso cuando las relaciones no son lineales o no siguen una distribución normal. Para mejorar la precisión y confiabilidad de los resultados, se emplea la técnica de bootstrap, generando múltiples muestras aleatorias de los datos y capturando su variabilidad. Con estas muestras, se realiza un análisis detallado de la correlación de Spearman entre las variables calculadas y la puntuación total del PHQ-8.

### 3. Resultados y Discusión

Primeramente, se analizan las variables calculadas según tres niveles de depresión (Nivel 0 si PHQ-8<10, N=797; Nivel 1 si 10≤PHQ-8<20, N=569; Nivel 2 si PHQ-8≥20, N=168), y tres tipos de umbrales para definir periodos de actividad, tal y como se muestra en la Tabla 1. Se pueden observar ciertas tendencias en los biomarcadores de HRR según el umbral de actividad física definido. Por ejemplo, se puede ver que el T95 aumenta en los tres grupos de depresión cuando se utiliza el umbral de 60 pasos/min para definir actividad física, mostrando que la recuperación promedio requiere más tiempo. Sin embargo, si comparamos los valores según el nivel de severidad de la depresión para cada uno de los grupos, no parece haber patrones o diferencias visibles en las medianas.

Para ver si realmente había diferencias significativas entre grupos, se realizó el test U de Mann-Whitney. Dado que no se observan grandes diferencias entre los grupos de depresión con puntuaciones PHQ-8 mayores de 10 y 20 (ver Tabla 1), se decidió agrupar ambos niveles para

obtener una muestra más homogénea a la hora de comprobar diferencias estadísticas. De esta manera, quedó un grupo de muestras de los valores cuando se habían recuperado de depresión (PHQ-8<10), y un grupo de muestras con depresión (PHQ-8 ≥ 10). En la Tabla 2 se presentan los intervalos de confianza del p-valor del test de U de Mann-Whitney habiendo utilizado Bootstrap, evaluando si existen diferencias significativas entre los grupos.

Variable y Severidad	Sin restricción	>60 pasos/min	Personalizado	
Capacidad [bpm]	0	8 (6, 12)	19 (14, 23)	21 (15, 26)
	1	9 (6, 12)	18 (14, 24)	21 (15, 27)
	2	8 (6, 11)	18 (14, 23)	21 (15, 26)
NW	0	177 (117, 242)	20(13, 30)	8(5, 13)
	1	154 (114, 218)	19 (13, 31)	9(6, 13)
	2	147 (103, 196)	18 (13, 25)	7(5, 11)
HR <sub>max</sub> [bpm]	0	92 (86, 100)	110 (103, 117)	81 (75, 92)
	1	106 (100, 112)	110(101, 117)	81(75, 92)
	2	107(100, 113)	110 (103, 115)	84 (77, 93)
HR <sub>min</sub> [bpm]	0	81 (75, 92)	87 (79, 95)	88 (81, 96)
	1	81 (75, 92)	86 (79, 95)	88(80, 96)
	2	84 (77, 93)	87 (82, 95)	87 (79, 97)
T95 [min]	0	-0.25 (-0.67, 0.08)	1.5 (0.75, 2.08)	1.83 (1, 2.58)
	1	-0.25 (-0.67, 0.17)	1.42 (0.67, 2)	1.75 (0.92, 2.42)
	2	-0.25 (-0.75, 0.17)	1.25 (0.75, 2.08)	1.71 (1, 2.42)

**Tabla 1.** Distribución de las variables propuestas estudiadas, para caracterizar la recuperación cardiaca. Se muestra en formato mediana (pct25%, pct75%), para los tres niveles de depresión según el PHQ-8 (Niveles 0, 1 y 2), separado según los tres umbrales para definir actividad física (6MWT).

Variable	Sin restricción	>60	Personalizado
Capacidad	(0.0652, 0.9778)	<b>(0.0099, 0.9524)</b>	<b>(0.0176, 0.9685)</b>
NW	(0.0914, 0.9813)	<b>(0.0462, 0.9705)</b>	<b>(0.0182, 0.9663)</b>
HR <sub>max</sub>	(0.1483, 0.9859)	(0.0705, 0.9825)	<b>(0.0435, 0.9738)</b>
HR <sub>min</sub>	(0.1534, 0.9886)	(0.1267, 0.9825)	(0.1159, 0.9852)
T95	(0.0681, 0.9812)	<b>(0.0393, 0.9744)</b>	(0.0641, 0.9785)
DVA	<b>(0.0077, 0.8563)</b>	<b>(0.0027, 0.5405)</b>	(0.1448, 0.9852)
HR <sub>1</sub>	(0.1545, 0.9826)	(0.1208, 0.9853)	(0.0913, 0.9805)
HR <sub>2</sub>	(0.1475, 0.9880)	(0.1307, 0.9866)	(0.0775, 0.9758)
HR <sub>3</sub>	(0.1466, 0.9819)	(0.1185, 0.9799)	(0.1065, 0.9866)
HRR <sub>1</sub>	(0.0687, 0.9839)	<b>(0.0082, 0.9525)</b>	<b>(0.0144, 0.9572)</b>
HRR <sub>2</sub>	(0.0688, 0.9800)	<b>(0.0189, 0.9632)</b>	<b>(0.0364, 0.9773)</b>
HRR <sub>3</sub>	(0.0703, 0.9779)	<b>(0.0121, 0.9591)</b>	<b>(0.0241, 0.9693)</b>

**Tabla 2.** Percentiles 5% y 95% de los p-valores para el test de U de Mann-Whitney, comparando cada una de las variables según los dos grupos de niveles de depresión (Nivel 0 vs. Nivel 1y2).

Las condiciones de más de 60 pasos por minuto y el percentil personalizado tienden a mostrar p-valores más bajos que el percentil sin restricción, lo que sugiere que estas restricciones mejoran la sensibilidad del test de U de Mann-Whitney para detectar diferencias entre los grupos de depresión. Esto indica que un nivel mínimo de actividad es un requerimiento para evaluar la respuesta del sistema nervioso autónomo y buscar diferencias en los niveles de depresión. Sin embargo, variables como HR<sub>min</sub> y HR en 1, 2 y 3 minutos no muestran p-valores significativos en ninguna condición, lo que sugiere que no son sensibles a los cambios en los niveles de depresión o no están directamente relacionadas con esta condición. Además, la

variabilidad en los p-valores obtenidos mediante el análisis Bootstrap refleja la complejidad y fluctuación en la relación entre las variables fisiológicas y la depresión. Esta variabilidad puede deberse a diferencias en las respuestas fisiológicas de los participantes y a cambios en su estado emocional durante el estudio. Adicionalmente a lo presentado en las Tablas 1 y 2, también se ha analizado la correlación entre el nivel de depresión y cada una de las variables. Sin embargo, tampoco se encontraron niveles de correlación significativos. Por lo tanto, es crucial interpretar estos resultados con cuidado y considerar que un análisis más detallado o con una mayor muestra podría revelar patrones más consistentes.

La principal limitación de este trabajo es que no hay disponible una muestra de participantes control, sin depresión, con la que comparar los resultados. Esta falta de comparación podría estar ocultando diferencias importantes en los resultados. También la dificultad para encontrar diferencias significativas basadas en el PHQ-8, un test subjetivo cuya precisión depende de la autoevaluación del paciente, lo que complica correlacionar sus resultados con datos fisiológicos objetivos. Dado que el PHQ-8 no siempre captura todas las facetas de la depresión, sugerimos utilizar los datos fisiológicos como complemento. Además, no se han segregado datos por edad, género, etc., una línea futura es analizar separando por grupos, ya que la edad y el género influyen en las respuestas fisiológicas y psicológicas de los individuos [9]. Otra limitación se encuentra en la diferencia en los tiempos de muestreo entre ambas series de datos. La serie de pasos está muestreada cada minuto, mientras que la serie de frecuencia cardíaca (HR) presenta un muestreo irregular, con una frecuencia típica de 5 segundos.

Estas limitaciones subrayan la necesidad de cautela al interpretar los resultados de este estudio. Aunque se han identificado algunas tendencias, las restricciones metodológicas y contextuales impiden establecer conclusiones definitivas sobre la relación entre la depresión, la actividad física y el HRR. Futuras investigaciones deberían considerar estos factores y trabajar en la generalización de los resultados para obtener una comprensión más clara y robusta de estos fenómenos.

#### 4. Conclusión

En este trabajo se ha estudiado la recuperación cardíaca en una selección de 217 pacientes del estudio internacional RADAR-MDD, que investiga el uso de dispositivos wearables para medir parámetros como el HR y su relación con la depresión. Se han utilizado datos de HR, pasos e indicadores PHQ-8 como referencia de la severidad de la Depresión. Los biomarcadores fisiológicos obtenidos durante el estudio, aunque no muestran diferencias significativas según el nivel de depresión, pueden ofrecer información valiosa que complementa a las evaluaciones clínicas tradicionales, proporcionando una visión más completa del estado del paciente.

#### Agradecimientos

Este trabajo ha sido parcialmente financiado por los proyectos TED2021-131106B-I00, PID2021-126734OB-C21 (Ministerio de Ciencia e Innovación y Fondo Social Europeo), España, el grupo BSICoS T39-23R (Gobierno de Aragón y Fondo Social Europeo). El proyecto RADAR-CNS ha recibido financiación de Innovative Medicines Initiative 2 Joint Undertaking mediante el acuerdo No 115902. Los cálculos fueron realizados por el ICTS NANBIOSIS (HPC de la Universidad de Zaragoza).

#### Referencias

- [1] Evelyn Bromet, Laura Helena Andrade, Irving Hwang, Nancy A Sampson, Jordi Alonso, Giovanni De Girolamo, Ron De Graaf, Koen Demyttenaere, Chiyi Hu, Noboru Iwata, et al. Cross-national epidemiology of dsm-iv major depressive episode. *BMC medicine*, 9:1–16, 2011.
- [2] K. Kroenke et al. The PHQ-8 as a measure of current depression in the general population. *Journal of Affective Disorders*, 114:163–173, 04 2009.
- [3] Carmen Schiweck, Deborah Piette, Daniel Berckmans, Stephan Claes, and Elske Vrieze. Heart rate and high frequency heart rate variability during stress as biomarker for clinical depression. a systematic review. *Psychological medicine*, 49(2):200–211, 2019.
- [4] Spencer L James, Degu Abate, Kalkidan Hassen Abate, Solomon M Abay, Cristiana Abbafati, Nooshin Abbasi, Hedayat Abbastabar, Foad Abd-Allah, Jemal Abdela, Ahmed Abdelalim, et al. Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990–2017: a systematic analysis for the global burden of disease study 2017. *The Lancet*, 392(10159):1789–1858, 2018.
- [5] RADAR-CNS: Remote Assessment of Disease and Relapse – Central Nervous System. <https://radar-cns.org/>
- [6] Daivaras Sokas, Birutė Paliakaite, Andrius Rapalis, Vaidotas Marozas, Raquel Bailón, and Andrius Petrenas. Detection of walk tests in free-living activities using a wrist-worn device. *Frontiers in physiology*, 12:706545, 2021
- [7] Aicko Y Schumann, Jan W Kantelhardt, Axel Bauer, and Georg Schmidt. Bivariate phase-rectified signal averaging. *Physica A: Statistical Mechanics and its Applications*, 387(21):5091–5100, 2008.
- [8] Tim Hesterberg. Bootstrap. *Wiley Interdisciplinary Reviews: Computational Statistics*, 3(6):497–526, 2011.
- [9] Soraya Seedat, Kate Margaret Scott, Matthias C Angermeyer, Patricia Berglund, Evelyn J Bromet, Traolach S Brugha, Koen Demyttenaere, Giovanni De Girolamo, Josep Maria Haro, Robert Jin, et al. Cross-national associations between gender and mental disorders in the world health organization world mental health surveys. *Archives of general psychiatry*, 66(7):785–795, 2009.