

RESUMEN GRADE DEL ENSAYO CLÍNICO:

Ensayo clínico de 6 semanas sobre un único tratamiento con Hipertermia del Cuerpo Entero para el trastorno depresivo mayor.

Janssen CW, Lowry CA, Mehl MR, Allen JJ, et al. Whole-Body Hyperthermia for the Treatment of Major Depressive Disorder: A Randomized Clinical Trial. *JAMA Psychiatry*. 2016 May 12;73(8):789-95.

Abreviaturas: **BDI-II:** Inventario de Depresión de Beck versión II; **DE:** desviación estándar; **DSM-IV:** Diagnostic and Statistical Manual of Mental Disorders (Manual Diagnóstico y Estadístico de trastornos mentales, 4ª edición); **ECA:** estudio controlado aleatorizado; **EDM:** episodio depresivo mayor; **HAM-D-17:** Escala de Depresión de Hamilton de 17 ítems; **HCE:** hipertermia del cuerpo entero; **TDM:** Trastorno Depresivo Mayor.

Nota sobre nuestros cálculos: Nosotros calculamos la p para la diferencia en las variables categóricas mediante chi cuadrado de Pearson, y mediante la adecuada diferencia de medias para variables continuas, con sus intervalos de confianza al 95% (IC 95%).

I. INTRODUCCIÓN.

Las vías sensitivas de la termorregulación se proyectan desde la piel hasta zonas específicas de la región cortical y subcortical, las cuales, están directamente implicadas en el registro y reacción a señales termales agradables. En pacientes con trastorno depresivo mayor (TDM) se observa que esas áreas cerebrales tienen menos actividad y además presentan anomalías en la termorregulación, caracterizadas por: aumento de la temperatura corporal y disminución de la capacidad de transpiración. Esas anomalías se normalizarían tras el tratamiento de Hipertermia del Cuerpo Entero (HCE).

Entre 2013 y 2015 Clemens y col llevaron a cabo un pequeño ensayo abierto en el que una sola sesión de HCE se asoció con una reducción en los síntomas depresivos en la primera semana post-intervención. Sin embargo, no pudo descontarse el posible efecto placebo derivado del aparataje y demás. Por esa razón diseñaron un ensayo clínico enmascarado para los pacientes mediante una simulación, a fin de averiguar si un tratamiento con una sola sesión de hipertermia del cuerpo entero puede inducir un efecto antidepresivo específico.

II. LO PROYECTADO.

A) OBJETIVO: Evaluar la reducción de síntomas en pacientes con TDM de un solo tratamiento con hipertermia de cuerpo entero (HCE) frente uno similar simulado sin hipertermia, y cuánto tiempo dura el efecto.

B) TIPO DE ESTUDIO: Estudio controlado, aleatorizado, y cegado para pacientes y evaluadores de los resultados, para comparar entre ellos sus respectivas reducciones entre antes y después en sus puntuaciones del DSM-IV, la Escala de Hamilton de Depresión (HAM-D) y demás cuestionarios diagnósticos. La comparación de los tres grupos se hace calculando la probabilidad p de que las diferencias sean debidas al azar, mediante la diferencia de medias y diferencia estandarizada de medias (d de Cohen, o tamaño del efecto) para las variables continuas. Las comparaciones entre el número de individuos que finalizan sin diagnóstico de TDM se hacen mediante chi cuadrado de Pearson.

Los autores no informan de la hipótesis de la que parten en la reducción de las puntuaciones de las escalas en el grupo de estándar de antidepresivos para poder calcular el tamaño de la muestra.

C) POBLACIÓN ESTUDIADA Y CRITERIOS DE INCLUSIÓN Y EXCLUSIÓN.

1º Criterios de inclusión: Voluntarios con 18-65 años que cumplen los criterios de TDM según el DSM-IV, y una puntuación ≥ 16 en la HAM-D, médicamente sanos y sin tomar psicofármacos.

2º Criterios de exclusión: Diagnóstico de claustrofobia, anorexia nerviosa o bulimia, de abuso de alcohol o sustancias, evidencia de riesgo de suicidio, demencia (eje I), cualquiera del eje II que se juzgue que puede no comprometerse al tratamiento y seguimiento. Uso de psicotrópicos en las dos semanas previas (8 si fluoxetina), con la excepción de hipnóticos. Psicoterapia en las 2 semanas previas. Contraindicaciones con el aparataje. Embarazo o lactancia.

D) VARIABLES A MEDIR.

1º Variable primaria: a) Nº de pacientes que alcanzan una puntuación < 7 en la HAM-D; b) Cambio en las puntuaciones en la HAM-D desde el inicio.

2º Variables secundarias: Cambio en las puntuaciones en: a) el Inventario de Sintomatología Depresiva autoinformado (Inventory of Depressive Symptomatology-Self Report, IDS-SR); b) la Escala de discapacidad de Sheehan (Sheehan Disability Scale); y c) la Cuestionario de calidad de vida, satisfacción y placer, versión corta (Quality of Life Enjoyment Satisfaction Scale —short-form).

III. LO CONSEGUIDO.

A) ASIGNACIÓN DE LOS SUJETOS A LOS GRUPOS.

1º ¿Se efectuó la aleatorización?: Sí, mediante un ordenador central a distancia, con bloques de 6 (3 y 3 de cada).

2º ¿Se mantuvo oculta la asignación de los grupos para los reclutadores?: Sí.

3º Participantes que fueron al grupo de intervención y de control.

- a) Grupo de hipertermia del cuerpo entero (HCE), 16 individuos.
- b) Grupo simulado, 14 individuos.

4º ¿Resultaron similares en el inicio los grupos de intervención y control con respecto a los factores pronósticos conocidos?: De las características sociodemográficas y clínicas de las que se tomaron datos, no hubo diferencias estadísticamente significativas, salvo en el Cuestionario de Credibilidad / Expectativa, tal como mostramos en la **tabla 1**.

Tabla 1: Características sociodemográficas y clínicas en el inicio

	Nº (%) o Media (DE)		valor de p
	Grupo de intervención, HCE; n= 17	Grupo de control, simulado; n=17	
Mujeres	12/17 (70,59%)	11/17 (64,71%)	0,714
Edad, años	36,71 (DE 15,2)	41,47 (DE 12,54)	0,327
Blancos no hispanos	9/17 (52,94%)	10/17 (58,82%)	0,730
HAM-D	20,71 (DE 4,87)	22,76 (DE 4,42)	0,208
Cuestionario de Credibilidad	0,89 (DE 2,72)	-1,01 (DE 2,5)	0,042
Cuestionario de Expectativa	1,02 (DE 2,68)	-1,26 (DE 1,91)	0,007
Duración del episodio actual, días	126,07 (DE 163,82)	100,5 (DE 143,67)	0,632
Episodios depresivos previos/ paciente	0,87 (DE 1,06)	0,5 (DE 0,76)	0,251
Uso previo de antidepresivos	9/17 (52,94%)	8/17 (47,06%)	0,732
Exposiciones a antidepresivos/ paciente	0,94 (DE 1,14)	0,76 (DE 1,09)	0,641
Uso previo de hipnóticos	3/17 (17,65%)	0/17 (0%)	0,070

HCE: Grupo de hipertermia del cuerpo entero; DE: desviación estándar.

NOTA: En el Cuestionario de Expectativa se consideran baja por debajo de cero y alta por encima.

5º ¿Se mantuvo oculta la asignación de los grupos para los investigadores que hacen el seguimiento?: No. **¿Y para los pacientes?:** No se les reveló., pero como en este caso esto no es suficiente para garantizarla, tras la finalización de la intervención se les pidió a todos los pacientes que adivinaran si habían recibido el tratamiento real o el simulado. Adivinaron a qué grupo pertenecían 15/16 (93,75%) del grupo de HCE y 4/14 (28,57%) del grupo simulado. Esto significa que el *efecto expectativa*, por estar o creer estar siendo tratados con la intervención HCE, pudo generarse en 15/16 (93,75%) y en 10/14 (71,43%) respectivamente, $p = 0,102$, aunque los primeros estaban acertados y los segundos equivocados.

6º ¿Se mantuvo oculta para los investigadores que hacen las evaluaciones mediante los cuestionarios diagnósticos y/o asignan los eventos?: Sí.

B) SEGUIMIENTO, ABANDONOS Y PÉRDIDAS.

1º Pauta de tratamientos y cuidados:

Antes del inicio todos los participantes respondieron al Cuestionario del historial de tratamiento con antidepresivos (del Hospital General de Massachusetts), y al Cuestionario de Credibilidad/ Expectativa.

Grupo de intervención, HCE. Durante 80 a 140 minutos, a cada individuo, en posición yacente, se le aplica calor con el sistema Heckel HT3000 hasta alcanzar una temperatura central de 38,5 °C, mediante la aplicación de calor al tronco con luces infrarrojas, y a las extremidades con bobinas de calentamiento infrarrojo. A continuación, se procede a una fase de enfriamiento de 60 minutos, permaneciendo en posición yacente. El dispositivo monitoriza temperatura central, periférica, presión sanguínea e hidratación.

Grupo de simulación. El procedimiento fue igual, excepto que utilizaba luces naranjas no calefactoras y un ventilador falso, para producir un color y ruido similares, pero sin calor infrarrojo en el tronco ni en las extremidades. De esta forma se alcanza una temperatura central de 37,7 °C.

A ambos grupos se le colocaron sensores para monitorizar constantemente las temperaturas central y periférica. La temperatura media periférica del grupo intervención fue de 40,7 °C y del grupo simulado de 39,7 °C.

2º Tiempo de seguimiento conseguido: 6 semanas.

3º ¿Se detuvo el estudio antes de lo proyectado?: No.

4º Abandonos y pérdidas.

Abandonos totales: Sin contar 1 vs 3 que no comenzaron a tratarse con la intervención ni el control, hubo 1 vs 3 que se trataron con sus respectivos tratamientos, pero que abandonaron el seguimiento. De éstos hubo 0/16 (0%) vs 2/14 (14,29%) por empeoramiento de los síntomas, $p=0,118$.

Pérdidas de seguimiento: No hubo ninguna en esta fase.

5º Se efectuó análisis por (intención de tratar, protocolo...): Por intención de tratar, considerando no a los 17 y 17 pacientes que fueron aleatorizados, sino a los 16 y 14 que recibieron el tratamiento.

C) RESULTADOS.

1º Magnitud y precisión de los resultados.

1. Tasas de individuos que alcanzan una puntuación ≤ 7 en la Escala de Hamilton de Depresión. Al cabo de 6 semanas se encontró una diferencia estadísticamente significativa entre los 0/14 (0%) en el grupo simulado frente a los 6/15 (40%) en el grupo de HCE, NNT 3 (1 a 6), como mostramos en la **tabla 2**. Los investigadores definen estipulativamente esa condición como "remisión".

	Nº (%), Grupo de control, Simulado	Nº (%), Grupo de intervención, HCE (**)	RR (IC 95%)	RAR (IC 95%)	NNT (IC 95%)	Valor de p
Puntuación ≤ 7 HAM-D, tras 6 semanas	0/14 (0%)	6/15 (40%)	-----	40% (15,75% a 69,51%)	3 (1 a 6)	0,008

(*) Los investigadores definen estipulativamente esa condición como "remisión". (**) HCE: hipertermia del cuerpo entero.

2. Cambios en la Escala Hamilton de Depresión (HAM-D): Al comparar los cambios entre las puntuaciones en el inicio y las puntuaciones en las sucesivas semanas, se encontraron diferencias estadísticamente significativas en las reducciones de la puntuación HAM-D en las semanas 1ª, 4ª y 6ª a favor del grupo de intervención, tal y como mostramos en la **tabla 3**.

	HAM-D, medias				
	Grupo de intervención, HCE (*)	Grupo de control, Simulado	Diferencia de los cambios (IC 95%)	d de Cohen (IC 95%)	valor de p
En el inicio, n= 16 vs 14	20,71 (DE 4,87)	22,75 (DE 4,42)			
Cambio entre la semana 1 y el inicio, n= 15 vs 14	-5,91	-1,89	-4,02 (-7,42 a -0,62)	-0,92 (-1,71 a -0,13)	0,022
Cambio entre la semana 2 y el inicio, n= 15 vs 12	-8,04	-4,04	-4 (-8,09 a 0,09)	-0,8 (-1,68 a 0,08)	0,055
Cambio entre la semana 4 y el inicio, n= 15 vs 11	-7,78	-4,96	-4,86 (-8,61 a -1,11)	-1,08 (-1,92 a -0,25)	0,013
Cambio entre la semana 6 y el inicio, n= 15 vs 11	-8,31	-5,54	-5,39 (-9,64 a -1,14)	-1,05 (-1,88 a -0,22)	0,015

(*) HCE: hipertermia del cuerpo entero.

Para comprobar si la diferencia estadísticamente significativa en el Cuestionario de Expectativa en el inicio podría haber sesgado los resultados de esta variable, los investigadores

hicieron un análisis de subgrupos (alta y baja), y tras practicar un test para calcular la probabilidad de una interacción, estimaron que el resultado de la cohorte completa era consistente, es decir que la diferencia en la puntuación de la expectativa no pareció afectar al resultado obtenido.

3. Cambios en las demás escalas y cuestionarios: Al comparar los cambios entre las puntuaciones en el inicio y las puntuaciones en la semana 6ª, no se encontraron diferencias estadísticamente significativas en las reducciones de la puntuación en: a) el Inventario de Sintomatología Depresiva; b) la Escala de discapacidad de Sheehan; y c) la Cuestionario de calidad de vida, satisfacción y placer, versión corta, tal y como mostramos en la **tabla 4**.

Tabla 4: Puntuaciones en las escalas y cuestionarios desde el inicio hasta los 6 meses de seguimiento, tras un único tratamiento.

	Medias				
	Grupo de intervención, HCE (*)	Grupo de control, Simulado	Diferencia de los cambios (IC 95%)	d de Cohen (IC 95%)	valor de p
Inventario de Sintomatología Depresiva autoinformado (Inventory of Depressive Symptomatology-Self Report, IDS-SR)					
En el inicio, n= 16 vs 14	40,2 (10,30)	39,93 (10,61)			
Cambio entre la semana 6 y el inicio; n= 15 vs 11	-16,33	-7,43	-8,9 (-1,21 a 19,01)	-0,71 (-1,52 a 0,1)	0,082
Escala de discapacidad de Sheehan (Sheehan Disability Scale)					
En el inicio, n= 16 vs 14	18,71 (5,53)	19,07 (5,02)			
Cambio entre la semana 6 y el inicio; n= 15 vs 11	-6,85	-3,5	-3,35 (-9,63 a 2,93)	-0,44 (-1,27 a 0,39)	0,282
Escala calidad de vida, disfrute y satisfacción, versión corta (Quality of Life Enjoyment Satisfaction Scale –short-form)					
En el inicio, n= 16 vs 14	34,67 (8,09)	36,93 (6,72)			
Cambio entre la semana 6 y el inicio; n= 15 vs 11	6,06	3,5	0,44 (-7,79 a 8,67)	0,04 (-0,79 a 0,88)	0,913

(*) HCE: hipertermia del cuerpo entero.

2º Efectos adversos: Dolor de cabeza, fatiga y sequedad de boca. En el grupo de intervención además aparecieron náuseas y sudoración. Los efectos adversos fueron leves y limitados en el tiempo, se habían resuelto 1 hora después del tratamiento. No se registraron efectos adversos graves.

D) COSTES: No se informa en el estudio ni aún siquiera de los costes directos.

IV. COMENTARIOS (DISCUSIÓN Y OPINIÓN DEL EVALUADOR).

Los investigadores señalan como limitación el pequeño tamaño de la muestra, lo que no aconseja hacer muchos análisis ajustando las covariables que no estaban equilibradas en el inicio, por el riesgo de cometer errores tipo 1 (dar por bueno algo que es malo).

Otra de las limitaciones que citan es la dificultad de descartar un efecto expectativa por el alto número de individuos que adivinaron estar en el grupo de intervención (93,7%), si bien nosotros hemos comprobado que hubo un 71,4% del grupo de simulación que creía haber sido tratado con la intervención, y por tanto generar el efecto expectativa del mismo modo que los anteriores.

V. CONFLICTOS DE INTERESES Y CALIDAD DEL ESTUDIO.

A) CONFLICTOS DE INTERESES. El estudio fue financiado por la Fundación para la Investigación del Cerebro y el Comportamiento (Premio Investigador Independiente), la Fundación para el Tratamiento Alternativo de los Trastornos Depresivos y Bipolares, el Instituto para la Investigación en Salud Mental, la Fundación Braun, y de Barry y Janet Lang, y de Arch y Laura Brown. Ninguno de los financiadores intervino en el diseño, recolección de datos, gestión, análisis y publicación de los datos.

De los 16 investigadores, 1 había recibido pagos de Merck. Ninguno recibió financiación de la empresa que fabrica el aparato Hecke HT300 de hipertermia d cuerpo entero. Ningún otro conflicto de intereses fue declarado.

B) CALIDAD DEL ESTUDIO (VALIDEZ DE LA EVIDENCIA).

¿Pregunta clara y precisa?: **Sí.**
¿Se efectuó una aleatorización correcta?: **Sí.**
¿Se mantuvo oculta la asignación de los grupos para los que hacen el reclutamiento?: **Sí. Se consiguió que creyeran estar en el grupo de intervención el 93,7% y el 71,4%, los primeros acertadamente y los segundos equivocadamente.**
¿Estaban equilibrados los factores pronósticos entre ambos grupos?: **Sí, salvo en el cuestionario expectativa.**
¿Se mantuvo oculta la secuencia de aleatorización para los investigadores que hacen el seguimiento?; **No.** ¿Y para los pacientes?; **Se consiguió que creyeran estar en el grupo de intervención el 93,7% y el 71,4%, los primeros acertadamente y los segundos equivocadamente.**
¿Y para los que asignan los eventos, y para los que obtienen los datos de laboratorio?: **Sí.**
¿Fue completo el seguimiento, cumpliendo con no detenerlo antes de lo previsto?: **Sí.**
¿Se contabilizaron los abandonos?: **Sí.** ¿Y las pérdidas?: **Sí.**
¿Se hicieron los cálculos por “intención de tratar”?: **Sí.**

Sistema GRADE: Calidad de la evidencia MODERADA. Justificamos la rebaja por: **1)** Tamaño de la muestra muy pequeño, que puede producir: a) desequilibrios no detectados en el inicio; y b) errores tipo 1 al final tras los ajustes por técnicas estadísticas. **2)** Tiempo corto. **3)** Si bien no parece haber dudas en la estimación de depresión grave y ausencia de depresión, los estadios intermedios no pueden ser determinados con una prueba objetiva, y aún menos aislados del carácter histórico de la cultura en la que el individuo está y de la que es. Por tanto, no es posible la validación del DSM-IV frente a una prueba objetiva porque ésta no existe, de modo que, aunque es conocido el riesgo de falsos positivos en los estadios intermedios, su porcentaje no puede cuantificarse.

V. CONCLUSIONES DEL ENSAYO CLÍNICO.

Con una calidad de evidencia moderada, **para individuos de 39 años (DE 13) con diagnóstico de trastorno depresivo mayor según el DSM-IV, y puntuación 22 (DE 4,5) en la HAM-D,** después de 6 semanas tras un único tratamiento: a) hay una diferencia estadísticamente significativa entre el 40% de individuos que alcanzan una puntuación ≤ 7 en el grupo de hipertermia del cuerpo entero (HCE) y el 0% en el grupo de control simulado; b) hay una diferencia estadísticamente significativa entre el cambio de -8,31 puntos en la HAMD-D desde el inicio en el grupo de HCE frente al cambio de -5,54 en el grupo de control simulado; c) no hay diferencias estadísticamente significativas en los cambios en el Inventario de Sintomatología Depresiva, Escala de discapacidad y Escala calidad de vida, disfrute y satisfacción; y d) no se registró ningún efecto adverso grave.

No se informa en el estudio de los costes, ni aún siquiera de los directos.

ADDENDA: BREVE DESCRIPCIÓN DE LOS INSTRUMENTOS DIAGNÓSTICOS

HAMD-17: Escala de Depresión de Hamilton de 17 ítems.

La Escala de Depresión de Hamilton (HAMD) es una escala heteroaplicada diseñada para evaluar la severidad de los síntomas en personas previamente diagnosticadas con TDM, y valorar los cambios en la sintomatología a lo largo del tiempo. Consta de 17 ítems que recogen los principales síntomas

cognitivos, somáticos, conductuales y vegetativos de la depresión, cada uno de ellos con una posible puntuación de 0-2 ó 0-4, siendo la puntuación total máxima 52 y la mínima 0.

Aunque en este estudio se consideró una depresión ligera puntuación de 13 a 18, y de moderada más de 18, los puntos de corte recomendados por la Guía Clínica NICE son los siguientes: 0-7: No deprimidos; 8-13: Depresión ligera/menor; 14-18: Depresión moderada; 19-22: Depresión severa; ≥ 23 : Depresión muy severa.

Validez: Su correlación con otros instrumentos de valoración de la depresión como la Escala de Depresión de Montgomery-Asberg, o el Inventario de Sintomatología Depresiva y la Escala de Melancolía de Beck, oscila entre 0,8 y 0,9. La validez es menor en pacientes de edad elevada, aunque ha mostrado buenos índices psicométricos en subpoblaciones de especiales características y mantiene un buen rendimiento en población geriátrica. Es muy sensible al cambio terapéutico.

Cuestionario de Credibilidad / Expectativa

La credibilidad se ha definido como “cuán creíble, convincente y lógico es el tratamiento”, mientras que la expectativa se refiere a “mejoras que los clientes creen que se lograrán”. Una creencia, sin embargo, contiene componentes tanto cognitivos como afectivos. Lo que una persona piensa lógicamente que es el caso puede diferir de lo que se cree que es el caso. Estudios previos han mostrado una correlación más alta con el resultado final en el factor “expectativa” que en el factor “creencia”.

Inventario de Sintomatología Depresiva autoinformado (Inventory of Depressive Symptomatology-Self Report, IDS-SR)

Consta de 30 ítems, de los que se contestan 28, con cuatro opciones cada uno, para que el paciente conteste desde el cero (lo que menos) a tres (lo que más) lo que mejor describe su situación en los últimos 7 días. Los 28-30 ítems evalúan todos los dominios de síntomas designados por la DSM-IV para el diagnóstico de un episodio depresivo mayor. Estas evaluaciones se pueden utilizar para la detección de la depresión, a pesar de que se han usado predominantemente como medidas de la gravedad de los síntomas. El período de siete días antes de la evaluación es el plazo habitual para evaluar la gravedad de los síntomas. Los puntos de corte que orientan en la graduación de los síntomas de depresión son: nada: 0-13 puntos; leve: 14-25; moderada: 26-38; severa: 39-48; muy severa: 49-84 puntos.

Escala de discapacidad de Sheehan (Sheehan Disability Scale)

Es un cuestionario de autopercepción de 5 ítems, diseñado para medir el grado en que la discapacidad de un paciente, debido a una enfermedad o problema de salud, interfiere con el trabajo / escuela, actividades de la vida / de ocio, sociales, vida familiar / hogar responsabilidades. En los tres primeros ítems, se pide a los encuestados que indiquen en una escala que va del 0 (nada) al 10 (mucho) en qué magnitud estiman que sus síntomas han interrumpido sus actividades regulares durante la última semana en cada una de las tres subescalas: incapacidad para el trabajo, discapacidad para la vida social, y discapacidad para la vida familiar. No hay un punto de corte canónico, pero en cada subescala más de 5 puntos sugiere deterioro, y éstas pueden evaluarse de forma independiente o combinarse para una calificación global de deterioro funcional, que van de 0 (poco) a 30 (mucho). Hay dos preguntas adicionales sobre el número de días en que sus síntomas: les impiden asistir a la escuela, o les causan bajo rendimiento.

Cuestionario de calidad de vida, satisfacción y placer, versión corta (Quality of Life Enjoyment Satisfaction Scale —short-form).

Consta de 16 ítems, de los que se contestan 14, con cinco opciones cada uno, para que el paciente conteste desde el uno (muy malo) a cinco (muy bueno) lo que mejor describe cuán fue su satisfacción en los últimos 7 días, de modo que la puntuación global va desde 14 (muy malo) a 70 (muy bueno). Los 14 ítems evalúan la autopercepción en: salud física, humor, trabajo, actividades en el hogar, relaciones sociales, relaciones familiares, actividades en el tiempo de ocio, funcionamiento en la vida diaria, interés y/o rendimiento sexual, estatus económico, situación de su vivienda, capacidad para moverse físicamente sin sentirse mareado, inestable o caerse, su visión en términos de capacidad de hacer trabajos o pasatiempos.