



COMILLAS
UNIVERSIDAD PONTIFICIA

ICAI

ICADE

CIHS

METAANÁLISIS:
USO DE LA INTELIGENCIA ARTIFICIAL EN
LA REDUCCIÓN DE SINTOMATOLOGÍA
DEPRESIVA Y ANSIOSA.

Autor: Carmen Lorenzo Martel

Tutor: Pedro María Ruiz de Assín-Varela

Facultad de Ciencias Humanas y Sociales
Madrid

Mayo de 2024

RESUMEN

El uso de las nuevas tecnologías ha adquirido un rol fundamental en entornos como el trabajo, ocio y entretenimiento de la población. Lo cual se ha visto aumentado tras la pandemia de la COVID-19 mostrando un impacto directo en el aumento de los dispositivos, además de generar sentimientos intensos de soledad y malestar entre la población. Atendiendo a la demanda social de hacer frente al malestar generado, en este metaanálisis se ha tratado de observar el posible uso de la inteligencia artificial, en especial de *chatbots*, para reducir sentimientos de depresión y/o ansiedad de entre la población. Para ello se han seleccionado un total de 14 artículos los cuales cumplen criterios de inclusión y exclusión, se ha realizado una revisión narrativa y además, un metaanálisis acerca de las correlaciones de ambas sintomatologías. Los resultados alcanzados muestran que el uso de *chatbots* puede derivar en una reducción de síntomas depresivos, pero no concluye una mejora en síntomas ansiosos.

Palabras clave: Inteligencia artificial, Chatbots, Depresión, Ansiedad, Metaanálisis.

ABSTRACT

The use of new technologies has become fundamental in various spheres of life such as work, leisure and entertainment. Which has been magnified after COVID-19 pandemic, which has directly contributed to the increase in device usage, while also exacerbating feelings of isolation and discomfort among population. In response to this social demand, this metaanalysis aimed to explore the potential utility of artificial intelligence, specifically chatbots being able to alleviate depression and/or anxiety among the population. For this purpose, 14 articles meeting specific criteria were selected, followed by a narrative review and metaanalysis of both symptoms. This findings suggest that the use of chatbots may lead to a reduction in depressive symptoms, although there is no evidence of improvement in anxious symptoms.

Key words: Artificial Intelligence, Chatbots, Depression, Anxiety, Metaanalysis.

INDICE

INTRODUCCIÓN	4
OBJETIVO	12
MÉTODO	13
Resultados de búsqueda	15
Variables	16
Análisis de datos	17
RESULTADOS	18
Características metodológicas y de la muestra.....	18
Tipos de programas de intervención aplicados	22
Impacto de las intervenciones	27
Resultado del metaanálisis sobre la depresión	28
Resultado del metaanálisis sobre la ansiedad.....	29
Análisis de moderadores.....	30
DISCUSIÓN	34
Conclusiones	39
Referencias	40
ANEXOS	50

INTRODUCCIÓN

La IA ha sido definida como la simulación de la inteligencia humana mediante la programación de software (Rivera y Sánchez, 2016).

En 1950 Alan Turing planteó la idea de que una máquina pudiese razonar o imitar el comportamiento humano (Turing, 2012). A través del “Test de Turing” buscaba evaluar si una máquina dotada con algoritmos adecuados podría simular la inteligencia humana, siendo imperceptible quién ideaba la respuesta, un software o un ser humano (Vidal, 2007). Posteriormente, John McCarthy (1956) acuñó el término “Inteligencia Artificial (IA)” para hacer referencia a dispositivos electrónicos con capacidad potencial de simular el pensamiento y libre albedrío del ser humano (Ahuja, 2019).

En la Teoría de Procesamiento de la Información de Mahoney (1974) hace referencia al ser humano como un procesador de datos que responde a estímulos internos y externos, almacenándolos en la memoria y siendo capaz de generar respuestas parecidas ante estímulos con cierta similitud (Benítez et al., 2014). En el mismo artículo se hace referencia a la esencia del aprendizaje humano, siendo la exposición frecuente a situaciones acompañado de refuerzos positivos o negativos, en base a las decisiones tomadas. Atendiendo a este concepto, los profesionales de la IA han ideado un software que trata de simular el proceso de aprendizaje del ser humano (Castellaro, 2011).

Para conocer el procesamiento de la información de un software se ha de hacer mención al concepto “aprendizaje automático”, en inglés “Machine Learning” (ML), siendo la capacidad de dotar a las máquinas de la habilidad de aprender patrones a partir de los datos recibidos (García Milá, 2023; Mitchell, 2006). A modo resumen, puede ser definido como un flujo que comienza con la entrada de datos (inputs) sometidos a un procesamiento por algoritmos de aprendizaje para extraer patrones y conocimientos, produciendo resultados o predicciones finales a la salida (outputs) (Shalev-Shwartz y Ben-David, 2014).

Más recientemente, en el mundo de la inteligencia artificial se han desarrollado dos descubrimientos interesantes, entre los que se encuentra el Procesamiento de Lenguaje Natural (PLN), siendo un software instaurado en las máquinas que dota a los dispositivos informáticos de la capacidad de comprender, interpretar y manipular el lenguaje humano (Chávez, 2020). Y, *chatbots*, siendo agentes de conversaciones digitales que emplean PLN y el aprendizaje constante para mantener conversaciones a través de texto y/o voz buscando simular el comportamiento humano en un diálogo en desarrollo (Suta et al., 2020). Pérez y Céspedes (2021) lo definen como un programa informático diseñado para llevar a cabo, de forma inteligente y natural, conversaciones con seres humanos. Buscando lograr que los sistemas ofrezcan respuestas satisfactorias en un diálogo oral o escrito, de manera que resulten indistinguibles de las respuestas humanas (Rodríguez et al., 2014). En la actualidad, los *chatbots* se han convertido en la herramienta más usada dentro de lo que ofrece la inteligencia artificial (Rudolph et al., 2023).

En el último año, ha habido un crecimiento exponencial en el uso de *bots*, como por ejemplo, Chat Generative Pre-trained Transformer (ChatGPT), siendo el más usado hasta el momento (Motlagh et al., 2023). El uso de la IA emerge como la revolución tecnológica del siglo XXI (Ruiz y Velásquez, 2022), generado cambios en la manera en que la población hace uso de las nuevas tecnología (George et al., 2023). El Gobierno de España (2023), en su plan de Estrategia Nacional de Inteligencia Artificial (ENIA) indica que el uso de inteligencia artificial se encuentra en un momento de transformación único, con la llegada de nuevos recursos y aplicaciones en diversos ámbitos. Numerosos estudio consideran que la inteligencia artificial ha encontrado su lugar en la rutina diaria de nuestra sociedad, empleándose tanto en tareas de la vida cotidiana como en la automatización del trabajo (Benbya et al., 2020).

Su carácter novedoso, pone de manifiesto tanto las posibilidades como los riesgos que presentan las nuevas tecnologías. En la actualidad, los programas informáticos cuentan con condiciones favorables en el procesamiento de información, siendo considerablemente más

eficientes en la adquisición de datos debido a su capacidad de operar 24 horas al día y no contar con limitaciones humanas como el descanso, ocio, etc (Ávila-Tomás et al., 2020). Aunque en la actualidad muestra ciertas limitaciones éticas (Astobiza, 2020).

En cuanto al ámbito de la salud mental, el uso de *chatbots* ofrece nuevas posibilidades en el sector. Sus plataformas independientes de tiempo y ubicación permite un amplio acceso al paciente (Lin et al., 2017) además de preservar su anonimato, lo cual reduce las preocupaciones respecto a la estigmatización de temas delicados a tratar (Schnyder et al., 2017). Con su uso se facilita el registro de información, la posible identificación de patrones y tendencias, además de poder realizar predicciones (Paz-y-Miño, 2023). Permite almacenar altas cantidades de información y correlacionarla de forma rápida y eficiente, lo cual dista enormemente de las capacidades del ser humano (Paz-y-Miño, 2023).

Los *bots* conversacionales se consideran herramientas útil para brindar atención a la salud mental de la población con acceso limitado a servicios médicos o para personas que tienen dificultades a la hora de expresar sus emociones a un ser humano (Chávez, 2020). El mundo de las nuevas tecnologías amplía el espectro de alternativas terapéuticas a ofrecer al paciente, permitiendo así adaptarnos a sus necesidades, características y preferencias (González, 2023).

A pesar de los beneficios mencionados previamente, la incorporación de la IA es de carácter reciente en la población, por lo que en la actualidad plantea desafíos que deben ser abordados cuidadosamente (Aparicio, 2023). Atendiendo al cumplimiento de cuestiones éticas que rigen en el ámbito de la salud, se ha de manejar la información personal teniendo en cuenta la ley protección de datos y confidencialidad del paciente (Romero et al., 2020), además de solicitar su consentimiento y mantenerles informados durante el proceso. Es frecuente observar desconfianza en el paciente hacia posibles sesgos, errores o malas interpretaciones de la IA (Taylor et al., 2020, p. 325). Y es que las nuevas tecnologías no quedan exentas de cometer errores y sesgos (Friedler et al., 2019, p. 1).

Más allá de las potencialidades y desafíos a los que se enfrentan las nuevas tecnologías en el momento actual, su reciente aparición hace surgir numerosas dudas acerca del futuro de los trabajadores en la sociedad. Una de las principales incógnitas es si la inteligencia artificial tal como la conocemos hasta el momento es capaz de reemplazar o acompañar a los profesionales (Selma, 2021). Existen opiniones significativamente divergentes entre autores, Pau García Milá (2023), referente en el mundo del emprendimiento online, considera que “los humanos no competimos con la inteligencia artificial, sino con otros humanos que hacen uso de esta y consiguen sacarle el máximo provecho”. Su despliegue ha generado notables avances en sectores diferenciados como en el ámbito empresarial, las ciencias sociales y de la salud, incluyendo la psicoterapia (Chávez, 2020). Esta andadura hacia una sociedad más tecnológica ha resultado en que profesionales de diversas disciplinas hayan de reconsiderar sus modelos tradicionales (Rivera y Sánchez, 2016).

Un claro ejemplo de los cambios que ha generado la incorporación de la IA en el ámbito de la salud, lo muestra la medicina. Donde su uso ha sido incorporado como sistema de apoyo para asistir a profesionales en el diagnóstico de enfermedades y en la toma de decisiones relacionadas con el tratamiento (Amann et al., 2020). En la actualidad, podemos hablar del enfoque “Medicina 4P”, siendo un modelo de atención médica que busca transformar el modelo médico convencional enfocado en curar enfermedades, a un enfoque orientado a la promoción de la salud y bienestar del individuo (Gutiérrez y Febles, 2020). Se busca pasar de una ciencia reactiva a una ciencia más predictiva, preventiva, participativa y personalizada (Hood y Friend, 2011). Tratando de alcanzar una detección temprana de síntomas y posible diagnóstico a través de la identificación de factores de riesgo. Estos mismos autores, hablan de ofrecer una intervención adaptada a la individualidad del paciente, a sus gustos y necesidades. Estos avances pueden ser aplicables a otras ciencias de la salud como la psicología, proporcionando a los

terapeutas información valiosa para mejorar el compromiso terapéutico y por ende, aumentar el éxito en la terapia (Ruiz y Velásquez, 2022).

Colocando la mirada en el ámbito de la psicoterapia, hay quienes no conciben un lugar para la IA dentro del sector. Los expertos señalan que las terapias que carecen de contacto visual, escucha activa y relaciones cercanas entre el paciente y el terapeuta tienden a ser menos exitosas (Ruiz y Cano, 2002). Lambert (1986, 1992) en investigaciones anteriores, estudió los factores que explican el cambio del paciente durante un proceso terapéutico, concluyendo que aproximadamente entre el 15% y el 30% de dicho cambio se debe a las contribuciones del terapeuta, pudiendo destacar el uso de técnicas como la alianza terapéutica y factores comunes. Además, estudios posteriores como los de Martínez (1999) y Safran y Segal (1994) sostienen que alrededor de un 45% del éxito terapéutico se explica por variables inespecíficas, entre las que se encuentra el vínculo terapéutico y la percepción del paciente sobre habilidades empáticas, capacidad de escucha, confianza y aceptación del terapeuta. Wampold e Imel (2021) corroboran las conclusiones extraídas en investigaciones pasadas y resumen la eficacia terapéutica en tres agrupaciones: efectos del terapeuta, efectos específicos y los generales a causa de factores comunes. Estos autores consideran que con frecuencia los resultados en el ámbito de la psicoterapia dependen más del aporte del terapeuta que del propio tratamiento, además, la alianza terapéutica muestra un papel importante en el éxito terapéutico. La alianza terapéutica se ha definido como la calidad y la fuerza de la relación colaborativa entre el cliente y el terapeuta en el proceso psicoterapéutico (Vargas Gallegos et al., 2015).

Autores como Fiske et al. (2019) destacan la utilización de la IA como una herramienta complementaria para el terapeuta. Pudiendo ser un instrumento de ayuda en la orientación del diagnóstico y enfoque de la terapia, predicciones de riesgo y la disminución de errores médicos, resultando en una mejora de la productividad (He et al., 2020). Aunque, su uso no exime al terapeuta de la necesidad de supervisar y llevar a cabo un trabajo paralelo. Ya que, como destaca

Balsategui (2019), en la práctica real de la psicoterapia se ha de tener en cuenta hechos objetivos, pero tratar de alcanzar la exclusión total de la subjetividad resulta utópico. Además, su reciente aparición en el sector deriva en que ciertas implicaciones éticas y sociales aun estén pendientes de delimitar, por lo que se sugiere su supervisión (Fiske et al., 2019).

Dada la innovación y las oportunidades que ofrece el uso de inteligencia artificial en distintos ámbitos resulta interesante explorar cómo podría ser aplicada en el ámbito de la salud mental. Y es que, a lo largo de los años se ha observado como los trastornos mentales y del cerebro han supuesto un coste económico elevado en España. Estudios como Ruiz-Rodríguez et al. (2017) alcanzaron estimaciones de alrededor de 46 millones de euros en trastornos mentales, y la mitad de estos costes eran destinados a trastornos depresivos, ansiosos y somatizaciones. Además, el número de profesionales especialistas en Psicología Clínica trabajando en el Sistema Nacional de Salud Española (SNS) muestra ser muy reducido como para atender a la demanda social actual (Duro Martínez, 2021). Por todo esto y atendiendo a las demandas sociales actuales, se valora el uso *chatbots* como forma de hacer frente a las necesidades actuales del propio individuo y de la sociedad en general.

Y es que, a pesar del gran número de sujetos que presentan síntomas depresivos y/o ansiosos, es tan solo un pequeño porcentaje el que recibe un tratamiento adecuado (Patel et al., 2010). Generalmente como resultado de las barreras en la accesibilidad, conveniencia, preocupación por la estigmatización y los costes asociados a la terapia (Lindhiem et al., 2015; Fitzpatrick, Darcy y Vierhile, 2017). Esto da como resultado el que los *chatbots* se presenten como una plataforma innovadora para proporcionar psicoterapia superando ciertas limitaciones de la psicoterapia convencional (Lin et al., 2017).

La sintomatología depresiva ha sido definida por la Organización Mundial de la Salud (OMS, s.f) como una prolongada sensación de tristeza y ausencia de interés o disfrute en actividades que solían considerarse como gratificantes o placenteras, pudiendo incluir tanto

síntomas físicos como psicológicos. Según los datos de la Encuesta Europea de Salud en España (EESA, 2020), aproximadamente el 12,47% de la población española experimenta sintomatología depresiva de diversos niveles de gravedad. Entre los que se puede diferenciar la sintomatología leve (8,46%), síntomas moderados (2,52%), moderadamente graves (1,19%) y síntomas graves (0,58%). Además, la prevalencia muestra ser el doble en mujeres, 16,32%, frente a un 8,94% en hombres, independientemente de la gravedad de los síntomas. Además, la sintomatología depresiva tiende a aumentar con la edad, siendo más pronunciados el incremento a partir de los 75 años de edad, con un 24,03% en personas entre los 75 y 84 años.

Respecto a la ansiedad, se define como una tendencia a experimentar sensación de miedo o preocupación excesiva frente a situaciones (OMS, 2023). En el año 2023, la Confederación Salud Mental España y la Fundación Mutua Madrileña ha llevado a cabo un informe en el que destaca que un 42,1% de la población española, cuatro de cada diez personas han sufrido una depresión a lo largo de su vida. Porcentaje que aumenta con la ansiedad, donde un 47,6% sufre ansiedad o pánico y un 36,9% destaca haber vivenciado una ansiedad prolongada en el tiempo (Confederación Salud Mental España y Fundación Mutua Madrileña, 2023).

Tras muchos años de investigación y estudio, la literatura científica ha hecho explícita una relación entre ansiedad y depresión, existiendo una alta comorbilidad entre ambas sintomatologías (Clark y Watson, 1995; Lambert et al., 2004). Distintos profesionales en la temática señalan que normalmente la sintomatología ansiosa antecede a episodios depresivos (Rohde et al., 1991), mostrando una dificultad en separar una sintomatología de otra.

En lo que respecta a la intervención, se ha validado la eficacia de la psicoterapia tradicional cara a cara en el abordaje de la depresión y ansiedad (Patel et al., 2010). La terapia cognitivo-conductual (TCC) y terapia de solución de problemas (TSP) muestran ser las más mencionadas (Weitz et al., 2018). La TCC considera que las cogniciones disfuncionales desencadenan y perpetúan las emociones angustiantes y los comportamientos desadaptativos (Beck, 1970). Por

lo que su abordaje se centra en modificar las cogniciones disfuncionales, con motivo de mejorar el malestar emocional y los patrones de conducta no adaptativos (Hofmann et al., 2012). Por otro lado, D'Zurilla y Goldfrid (1971) hablan de la TSP, donde su intervención trata de fomentar una evaluación individualizada del problema e impulsar al individuo a identificar soluciones efectivas para afrontarlo. Trabajando la percepción de gravedad del problema y motivándole así a desarrollar estrategias diversas para contrarrestar y fortalecer su locus de control interno (Warmerdam et al., 2010).

Estos modelos de intervención han mostrado su eficacia a través de intervenciones cara a cara. De manera que los programas de intervención online tratan de mantener una estructura similar a la presentada en las intervenciones anteriores, y valorando los avances sociales y necesidades actuales, se presenta el uso de *chatbots* como un recurso esperanzador en el ámbito de la salud.

En los últimos años, a causa de la pandemia global de la SARS-CoV-2, más conocido como COVID-19, se ha observado un aumento tanto en la frecuencia como en la gravedad de los trastornos y sintomatología de salud mental (OMS, 2020; Vizard et al., 2020). Además, entidades como la Asociación Madrileña de Salud Mental han registrado un crecimiento en la necesidad de servicios de salud (AMSM, 2021). Esta crisis mundial ha modificado hábitos y rutinas en la población general, abandonando espacios de socialización para pasar a conectar de manera online (Aretio, 2021). Este mismo autor hace alusión a modificaciones de emergencia a modo de adaptación, como puede ser: el lugar de trabajo (pasando de la presencialidad al teletrabajo), los cambios en la digitalización de los cursos académicos, en la conciliación familiar y reparto de tareas, entre tantos otros. Estas transformaciones traen consigo un impacto en la salud mental, ya que como cita Aristóteles (384-322, a. de C.) “el hombre es un ser social por naturaleza” (Seidmann, 2000).

A lo largo de este tiempo, gran parte de la población pudo experimentar sentimientos de temor e incertidumbre ante el porvenir (Usher et al., 2020). La desinformación y ansiedad pública generó una gran repercusión en la estadística de la salud mental (Cedeño et al., 2020). Informes llevados a cabo por la Organización Mundial de la Salud (OMS, 2020) mostraban tanto un aumento en casos de depresión (27,6%) como de ansiedad (26,6%). Resultando en un aumento global de alrededor del 25% en la incidencia de trastornos de ansiedad y depresión a lo largo del primer año de pandemia (OMS, 2022). Entre los que se observa una mayor afectación en mujeres y personas jóvenes, especialmente en edades comprendidas entre los 20 y 24 años (OMS, 2020).

Como forma de concluir, a lo largo de los años y en especial tras la COVID-19 la población española ha vivenciado grandes cambios en el abordaje de situaciones. Evolucionando a una población más digitalizada, lo cual trae consigo sentimientos diversos entre la población pero en especial han destacado sentimientos de incertidumbre y malestar. Es por todos estos cambios sociales y teniendo en cuenta las demandas, gustos y posibles necesidades que muestran los sujetos en el momento actual, que se considera de especial interés valorar el lugar que ocupan las nuevas tecnologías en la sociedad. Y, centrando la mirada en el ámbito de la psicoterapia, el lugar que ocupa el uso de *chatbots*. Siendo esta la plataforma analizada a lo largo de este estudio, ya que su funcionamiento simula una conversación humana. Y por ende, se asemeja al estilo de comunicación de un proceso terapéutico.

OBJETIVO

El objetivo general de este trabajo es estudiar la eficacia de los programas de inteligencia artificial, centrándonos exclusivamente en el uso de *chatbots* para la reducción de sintomatología ansiosa y/o depresiva. Como objetivos específicos se plantea analizar si:

- Se da un cambio en el tamaño del efecto de síntomas depresivos y/o ansiosos tras recibir intervención a través de *chatbots*.

- Los cambios en la sintomatología del paciente son duraderos en el tiempo.
- La reducción de sintomatología se da tras la intervención exclusiva de la inteligencia artificial o hablamos de una intervención en paralelo (terapia ordinaria junto con inteligencia artificial).

Dar respuestas a estos objetivos nos permite alcanzar una visión global acerca de qué información disponemos en la actualidad y en qué medida se puede hacer uso de las nuevas tecnologías para abordar la sintomatologías mencionadas anteriormente.

MÉTODO

Estrategia de búsqueda

A lo largo del mes de noviembre de 2023 se realizó una búsqueda bibliográfica en las bases de datos Pubmed, Psycinfo y Psychology & Behavioral Sciences Collection (P&BSC). En la ecuación de búsqueda se hizo uso de los términos “*Chatbots*”, “*Anxiety*”, “*Depression*”, “*Randomized controlled trials*”, “*Psychotherapy*” y “*Artificial intelligence*”, etc. Se han empleado operador booleano como “AND” y “OR” entre cada una de las variables introducidas (**Tabla 1**). Para la búsqueda se ha hecho uso de lenguaje documental en todas sus variables exceptuando “*Artificial intelligence*”, “*Chatbots*”, “*Deprexis*” y “*Woebot*” que se introdujeron en lenguaje libre. El objetivo de esta búsqueda ha sido recopilar de manera amplia la información disponible hasta el momento sobre la temática a revisar.

Tabla 1*Ecuaciones de búsqueda en las bases de datos*

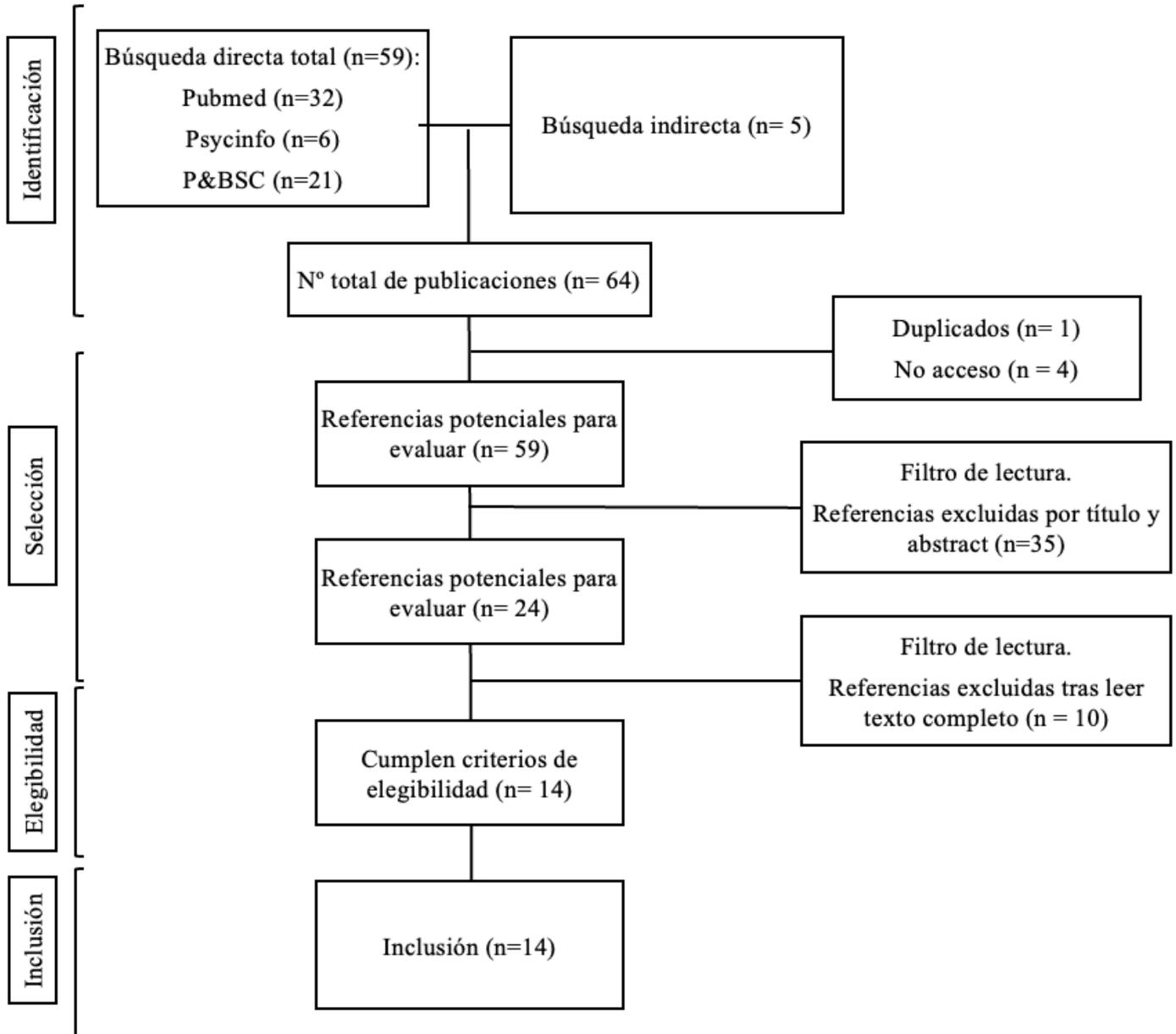
Base de datos	Nº referencias encontradas	Ecuación de búsqueda
Pubmed	32	("Depression"[Mesh] OR "Depressive Disorder"[Mesh] OR "Anxiety"[Mesh] OR "Anxiety Disorders"[Mesh]) AND ("Artificial intelligence" OR "chatbot" OR "Deprexis" OR "Woebot") AND "Psychotherapy"[Mesh] AND "Randomized Controlled Trial" [Publication Type]
PsycInfo	6	(DE "Depression (Emotion)" OR DE "Anxiety" OR DE "Anxiety Disorders") AND (DE "Chatbots" OR "Deprexis" OR "Woebot") AND DE "Randomized Controlled Trials" AND DE "Psychotherapy"
P&BSC	21	(DE "ANXIETY disorders" OR DE "ANXIETY" OR DE "MENTAL depression") AND (DE "CHATBOTS" OR DE "CHATGPT" OR "Online therapy" OR "Internet based treatment") AND DE "RANDOMIZED controlled trials" AND DE "PSYCHOTHERAPY"
Total	59	

Como criterios de elegibilidad se han considerado: (1) estudios experimentales – ensayos clínicos aleatorizados (ECA) (2) sujetos mayor de edad (3) sintomatología depresiva y/o ansiosa (4) población clínica y no clínica (5) uso de *chatbots* autónomos con intercambio bidireccional con el sujeto a través de dispositivos móviles o informáticos (6) publicados en español o inglés (7) año de publicación de los estudios entre 2009 y 2023.

Los criterios de exclusión considerados: (1) sujetos con una historia clínica ansiosa o depresiva grave y/o cronificada en el tiempo (2) individuos con intentos y/o ideación suicida (3) estudios no experimentales (4) sujetos menores de edad (5) población que no se encuentra en tratamiento (6) intervención exclusiva cara a cara (7) publicaciones en idiomas que no sean inglés o español (8) publicaciones posteriores al año 2009

Figura 1

Diagrama de flujo acerca de la selección de artículos



Resultados de búsqueda

Tras introducir la ecuación de búsqueda en las bases de datos mencionadas anteriormente, se identificaron 59 publicaciones. Además, se han añadido 5 publicaciones al análisis tras llevar a cabo una búsqueda indirecta, resultando en un total de 64 artículos. Entre estos, una de las publicaciones se mostraba duplicada y cuatro de ellas no permitían su acceso, siendo todas estas

excluidas desde un primer momento. Se realizó un primer cribado a través de la lectura de los títulos y de sus resúmenes, excluyendo aquellos que no cumplieren criterios de inclusión mencionados en este estudio lo cual redujo el número inicial de publicaciones a 24. Posteriormente, se procedió a una lectura exhaustiva de los artículos restantes, garantizando en todo momento el cumplimiento de los criterios de inclusión y exclusión. Una vez llevado a cabo este último cribado, el número final de artículos seleccionados para este análisis fueron 14 publicaciones (**Figura 1**).

Variables

Para comenzar, se han registrado variables metodológicas que guardan relación con la validez y fiabilidad de los estudios. Entre estas, incluye la evaluación del tipo de diseño empleado, el tamaño muestral, el tipo de medidas de resultado, el seguimiento empleado, la presencia de cegamiento tanto en los evaluadores como en los evaluados y el análisis de las pérdidas de sujetos a lo largo del estudio.

Adicionalmente, se han analizado variables vinculadas a las características de la muestra, pudiendo influir en la generalización de los resultados. Estas variables abarcan el rango de edad, la proporción de sexos, país de origen y el ámbito de reclutamiento (clínico o no clínico).

Tras haber considerado las características de la muestra, se ha examinado el tipo de intervención implementada, su extensión temporal, la variable evaluada, y los cuestionarios para la medida de variables.

Por último, se han evaluado las variables de resultado, dando respuesta al objetivo del estudio de analizar el posible cambio que produce el uso de *chatbots* en sintomatología depresiva y/o ansiosa. Se ha analizado si los cambios producidos son duraderos en el tiempo, además de observar si se debe exclusivamente al uso de las nuevas tecnologías o, si su efecto es debido a una intervención en paralelo con el tratamiento habitual.

Análisis de datos

Para llevar a cabo el análisis de los datos acerca de la eficacia de las intervenciones en los artículos seleccionados, se ha evaluado si la intervención ha alcanzado resultados de mejora, mantenimiento o empeoramiento de la síntomas de depresión y/o ansiedad.

Para cada estudios se calculó, en primer lugar, la diferencia de medias tipificadas para cada grupo (experimental y control), dividiendo la diferencia entre las puntuaciones post y pre-intervención por la desviación típica del momento pre. En segundo lugar, se calculó la ganancia del grupo experimental respecto del control, como la diferencia entre el cambio tipificado del grupo experimental menos el del grupo control (Botella y Sánchez-Meca, 2015). Las diferencias de medias fueron corregidas mediante la corrección propuesta por Hedges (Botella y Caperos, 2019). Por último, se calculó la varianza de dicha estimación de la ganancia siguiendo a Botella y Sánchez-Meca (2015, pp. 38-39).

El metaanálisis se ha hecho a través del programa Jasp (0.18.3), y las estimaciones se realizaron mediante máxima verosimilitud y con un modelo de efectos aleatorios. donde se ha evaluado la heterogeneidad de los estudios a través de los estadísticos Q (Huedo-Medina et al., 2006) e I^2 (Higgins y Thompson, 2002). Respecto al análisis de los resultados de depresión, cuando Q muestra ser significativo ($p < 0.01$) o $I^2 > 75\%$ se ha llevado a cabo un análisis de los moderadores, considerando la influencia en los datos del grupo control (LDE) o experimental (TH). En cuanto a la interpretación de los datos, se han seguido las guías propuestas por Cohen (1988): efecto pequeño (0.2), moderado (0.5) y grande (0.8).

RESULTADOS

Características metodológicas y de la muestra

Para la realización de este análisis se ha seleccionado un total de 14 artículos, entre ellos, seis recogen datos relacionados con sintomatología depresiva (Berger et al., 2011; Gold et al., 2023; Meyer et al., 2009; Moritz et al., 2012; Sandoval et al., 2017; Schroder et al., 2014), y los ocho restantes hacen referencia a ambas sintomatologías destacadas a lo largo de este análisis (Berger et al., 2017; Berger et al., 2018; Fitzpatrick et al., 2017; Fitzsimmons-Craft et al., 2021; Oh et al., 2020; Richards et al., 2015; Sadeh et al., 2023; Zwerenz et al., 2017).

Todas las investigaciones incluidas cuentan con un diseño experimental aleatorizado (ECA) entre los cuales once componen dos grupos (grupo experimental y control) (Berger et al., 2017; Berger et al., 2018; Fitzpatrick et al., 2017; Meyer et al., 2009; Moritz et al., 2012; Oh et al., 2020; Richards et al., 2015; Sadeh et al., 2023; Sandoval et al., 2017; Schroder et al., 2014; Zwerenz et al., 2017), los tres estudios restantes incluyen más de un grupo experimental (Berger et al., 2011; Fitzsimmons-Craft et al., 2021; Gold et al., 2023). Dentro de los grupos control seis de ellos tenían la condición de lista de espera (Berger et al., 2011; Fitzsimmons-Craft et al., 2021; Meyer et al., 2009; Moritz et al., 2012; Richards et al., 2015; Schroder et al., 2014), a la vez que otros siete estudios hablan de continuar con el tratamiento habitual, más reconocido en los estudios como TAU (Treatment as usual) (Berger et al., 2017; Berger et al., 2018; Fitzpatrick et al., 2017; Gold et al., 2023; Oh et al., 2020; Sadeh et al., 2023; Zwerenz et al., 2017). Tan solo en uno de los estudios se hace alusión a una condición de no tratamiento (NT) (Sandoval et al., 2017) (**Tabla A1**).

Entre los artículos con una condición experimental de tratamiento habitual (TH) podemos observar diferencias en base al tipo de intervención: artículos como Berger et al., (2017), Berger et al., (2018), Gold et al., (2023) y Sadeh et al., (2023) hacen referencia a una intervención psicoterapéutica “típica”, siendo sesiones semanales, cara a cara y de una hora de duración. Por

otro lado, hay grupos controles en los que tan solo se ofrece bibliografía al participante: Oh et al., (2020) ofrece un libro sobre trastornos de pánico y Fitzpatrick et al. (2017) el NIMH ebook (National Institute of Mental Health). Por último, Zwerenz et al., (2017) da información online semanal sobre la depresión, además de brindar al usuario con la posibilidad de recibir psicoterapia psicodinámica (**Tabla 3**).

En cuanto a las técnicas de control de variables usadas en los estudios, en gran parte de los ellos se realiza un análisis por intención de tratar (Berger et al., 2011; Berger et al., 2017; Berger et al., 2018; Fitzpatrick et al., 2017; Fitzsimmons-Craft et al., 2021; Moritz et al., 2012; Sadeh et al., 2023; Sandoval et al., 2017; Schroder et al., 2014; Richards et al., 2015), cegamiento (Gold et al., 2023; Schroder et al., 2014), enmascaramiento de datos (Gold et al., 2023; Meyer et al., 2009), bloqueo (Zwerenz et al., 2017) y, análisis de pérdida de sujetos (Berger et al., 2011; Berger et al., 2018; Sandoval et al., 2017; Schroder et al., 2014).

Para medir los síntomas de interés en este estudio se ha hecho uso de cuestionarios, a la hora de medir depresión uno de los cuestionarios más usados en la recogida de datos es el *Inventario de Depresión de Beck (BDI-II)*, siendo la medida seleccionada por diez de los estudios analizados (Berger et al., 2011; Berger et al., 2017; Berger et al., 2018; Gold et al., 2023; Meyer et al., 2009; Moritz et al., 2012; Richards et al., 2015; Sandoval et al., 2017; Schroder et al., 2014; Zwerenz et al., 2017). En la misma línea, se hace uso del *Patient Health Questionnaire (PHQ-8 o PHQ-9)* para recoger síntomas depresivos basados en criterios del DSM (Fitzpatrick et al., 2017; Fitzsimmons-Craft et al., 2021; Sadeh et al., 2023). Una escala que aúna ambos síntomas estudiados en este análisis podemos destacar la *Escala de Ansiedad y Depresión Hospitalaria (HADS)* usada en el estudio de Oh et al., 2020. En el caso de la ansiedad, la más destacada es la *escala de Trastorno de Ansiedad Generalizada (GAD-7)* usada por seis de los estudios analizados (Berger et al., 2018; Fitzpatrick et al., 2017; Fitzsimmons-Craft et al., 2021; Richards et al., 2015; Sadeh et al., 2023; Zwerenz et al., 2017) (**Tabla 3**).

Centrando la mirada en el tipo de intervención llevado a cabo, podemos identificar diferentes temporalidades, 2 semanas (Fitzpatrick et al., 2017), 4 semanas (Oh et al., 2020; Sandoval et al., 2017), 8 semanas (Fitzsimmons-Craft et al., 2021; Moritz et al., 2012; Richards et al., 2015; Sadeh et al., 2023), 9 semanas (Berger et al., 2017; Meyer et al., 2009; Schroder et al., 2014), 10 semanas (Berger et al., 2011) e incluso de 12 semanas (Berger et al., 2018; Gold et al., 2023; Zwerenz et al., 2017).

En cuanto al tipo de programa informático, el más compartido entre los autores destaca ser el Deprexis usado en estudios como Berger et al., (2011); Berger et al., (2018); Meyer et al., (2009); Moritz et al., (2012); Schroder et al., (2014) y Zwerenz et al., (2017). En el resto de programas señalados a continuación es tan solo uno de los autores quien hace uso de él, Woebot (Fitzpatrick et al., 2017), Velibra (Berger et al., 2017), Tessa (Fitzsimmons-craft et al., 2021), Eleos health platform (Sadeh et al., 2023), imbPST (Sandoval et al., 2017), y por último Amiria iCBT (Gold et al., 2023). Los distintos programas mencionados anteriormente serán descritos en profundidad más adelante.

Tras el proceso de intervención algunos estudios han llevado a cabo un seguimiento de los resultados, tratando de analizar si estos han sido mantenidos en el tiempo. Se realizó un seguimiento a los 2 meses (Sandoval et al., 2017), 3 meses (Gold et al., 2023) pero especialmente a los 6 meses (Berger et al., 2011; Berger et al., 2017; Berger et al., 2018; Meyer et al., 2009; Schroder et al., 2014). Aunque, tanto Fitzsimmons-craft et al., 2021 como Richards et al., 2015, llevaron a cabo seguimiento tanto a los 3 como a los 6 meses (**Tabla A1**).

En cuanto a la muestra de los estudios científicos, ha rondado entre los 41 y 700 participantes. Generalmente compuesto por población adulta, con un rango de edad amplio entre los 18 y 50 años. La media de edad entre todos los participantes se encuentra en 36,6 años. Todos los estudios muestran una proporción de mujeres mayor que de hombres, por lo general, la

distribución es de 70:30. Aunque existen excepciones, como es el caso de Oh et al., (2020) donde se observa una distribución similar entre hombres (48.78%) y mujeres (51.22%) (**Tabla A2**).

Teniendo en cuenta el tipo de participantes seleccionados para los estudios, podemos identificar población clínica, específicamente pacientes ambulatorios (Berger et al., 2011; Berger et al., 2017; Berger et al., 2018; Fitzsimmons-craft et al., 2021; Gold et al., 2023; Oh et al., 2020; Moritz et al., 2012; Sadeh et al., 2023; Sandoval et al., 2017; Schroder et al., 2014; Zwerenz et al., 2017). Y, población no clínica (Fitzpatrick et al., 2017; Meyer et al., 2009; Richards et al., 2015), siendo participantes los cuales consideraban bajo su propio criterio el tener síntomas depresivos y/o ansiosos (**Tabla A2**).

En términos generales, la selección de participantes atiende al criterio de inclusión de poseer síntomas depresivos y/o ansiosos. Aunque resulta de interés hacer mención a algunos de los diagnósticos recogidos, Berger et al., (2011) realiza su estudio con participantes diagnosticados de depresión mayor o distimia. Por otro lado, Berger et al., (2017) hace referencia a la ansiedad social, trastorno de pánico con o sin agorafobia y ansiedad generalizada, Fitzsimmons-craft et al., (2021) habla de una población con diagnóstico de trastorno de alimentación. Y finalmente tanto Gold et al., (2023) como Schroder et al., (2014) trabajan con una población con diagnóstico de esclerosis múltiple.

Entre los estudios seleccionados se destacan distintas variables que pudiesen estar incidiendo sobre los resultados alcanzados. Se ha de valorar el uso de medicación en el momento actual de la investigación, en el caso de Richards et al., (2015) contabilizan un 20,7% a diferencia de Zwerenz et al., (2017) donde cuentan con un 65,2% de los participantes. Aunando los estudios que lo analizan, una media de 36,25% de los participantes se encontraban haciendo uso de medicación en ese momento. Por otro lado, se recogen variables como el hecho de haber asistido previamente a un tratamiento psicológico (Berger et al., 2011; Berger et al., 2017; Berger et al., 2018; Gold et al., 2023; Meyer et al., 2009; Moritz et al., 2012; Richards et al., 2015; Zwerenz

et al., 2017), entre un 50% - 70% de los sujetos habían mantenido un tratamiento psicológico previo, con una media del 63%. Autores como Berger et al., (2017), Meyer et al., (2009) y Moritz et al., (2012) recogían el que un paciente se encontrase en tratamiento psicológico en el momento actual del estudio.

Respecto a la muestra, podemos observar distintos países de intervención entre los que predomina Alemania (Berger et al., 2011; Berger et al., 2017; Berger et al., 2018; Gold et al., 2023; Meyer et al., 2009; Moritz et al., 2012; Schroder et al., 2014; Zwerenz et al., 2017) seguido por Estados Unidos (Fitzpatrick et al., 2017; Fitzsimmons-Craft et al., 2021; Gold et al., 2023; Sadeh et al., 2023; Sandoval et al., 2017). Aunque también podemos encontrarnos con estudios llevados a cabo en países como Austria (Berger et al., 2017), Suiza (Berger et al., 2011; Berger et al., 2017), Corea del sur (Oh et al., 2020) e Irlanda (Richards et al., 2015).

El reclutamiento generalmente se llevó a cabo a través de anuncios en plataformas de internet como pueden ser redes sociales, páginas webs (Berger et al., 2011; Berger et al., 2017; Fitzpatrick et al., 2017; Fitzsimmons-Craft et al., 2021; Gold et al., 2023; Meyer et al., 2009; Moritz et al., 2012; Sadeh et al., 2023; Sandoval et al., 2017; Schroder et al., 2014; Zwerenz et al., 2017). Aunque también ha habido estudios que han reclutado a participantes tras intervenciones cara a cara con ellos (Berger et al., 2017; Berger et al., 2018; Oh et al., 2020; Richards et al., 2015) (**Tabla A2**).

Tipos de programas de intervención aplicados

Entre los 14 estudios seleccionados podemos encontrarnos con distintos tipos de programas informáticos que trabajan en busca de la reducción de sintomatología depresiva y/o ansiosa. Analizando el tipo de intervención llevado a cabo, el *Deprexis* ha sido el más usado por los autores (Berger et al., 2018; Meyer et al., 2009; Moritz et al., 2012; Schroder et al., 2014; Zwerenz et al., 2017). Destaca ser un programa con especial mirada sobre la práctica activa en el tratamiento de la depresión, en él se proporciona información psicoeducativa y ejercicios

basados en una intervención cognitivo-conductual. Buscando alcanzar el objetivo principal de reducir los síntomas depresivos. Este modelo de *chatbot* contiene diez módulos de contenido y uno de resumen distribuidos de la siguiente forma:

(1) Introducción.

Psicoeducación sobre depresión y se exploran necesidades y preferencias del paciente.

(2) Activación conductual y trabajo de las necesidades psicológicas básicas.

En la intervención con depresión se ha demostrado la mejora al trabajar con técnicas de reestructuración cognitiva y activación conductual (Cuijpers et al., 2007; Hollon et al., 2002). En ellas se anima a los usuarios a programar actividades que puedan satisfacer sus necesidades psicológicas básicas como las relaciones sociales, autonomía, competencias, autoestima y el disfrute hedónico (Lejuez et al., 2001).

(3) Reestructuración cognitiva.

Se da especial importancia a los pensamientos automáticos y su impacto en el estado anímico, se ofrecen técnicas sencillas para cuestionar, refutar, aceptar y tratar de dar distancia a esos pensamientos (Hautzinger et al., 2003).

(4) Mindfulness y aceptación.

Trabajo en la aceptación de experiencias no deseadas a través de la terapia de aceptación y compromiso (ACT) y la terapia cognitiva del mindfulness (MBCT). Sus hallazgos aportan grandes avances a la terapia tradicional (Hayes et al., 1999; Segal et al., 2002).

(5) Habilidades interpersonales.

Inciendo sobre las dificultades en el ajuste interpersonal en pacientes depresivos se trabaja la mejora en la comunicación verbal, no verbal y pautas de comportamiento como forma de mejorar las relaciones (Coyne, 1976; Joiner, 2000; Gable et al., 2004).

(6) Ejercicios de relajación, actividad física y cambios en su estilo de vida.

El ejercicio físico, un estilo de vida saludable y el uso de la relajación tiene un impacto directo sobre la sintomatología (Dunn et al., 2005; Richardson et al., 2005).

(7) Resolución de problemas.

Se ofrecen herramientas para aprender a definir los problemas, evaluar las opciones de solución, poner en práctica la solución elegida y finalmente evaluar los resultados con respecto al problema inicial.

(8) Experiencias infantiles y tempranas.

La literatura científica incide sobre la relevancia de las experiencias infantiles problemáticas en etapas posteriores de la vida (Addis et al., 1995; Harkness y Monroe, 2002).

(9) Psicología positiva e intervenciones centradas en la emoción.

Autores como Seligman et al., (2006) concluían que las intervenciones desde la psicología positiva como animar a las personas y expresar gratitud, pueden incidir en la reducción de síntomas depresivos.

(10) Trabajo sobre los sueños e intervenciones centradas en la emoción.

A pesar de no ser un elemento característico a trabajar dentro del mundo de la terapia cognitivo-conductual, se ha considerado relevante ofrecer un lugar a los sueños dentro de esta intervención (Hill et al., 2006). Se ofrecen técnicas básicas de cómo llevar un diario de sueños y reescribir los sueños cargados de problemas con motivo de darles un final positivo (Morris, 1985).

Siguiendo una línea similar a la estructura presentada anteriormente, Gold et al. (2023) hablan de Amiria, un programa de intervención que trata la esclerosis múltiples y la sintomatología que lo envuelve.

Otro de los programas usado por Berger et al. (2017) fue Velibra, un programa de intervención online que cuenta con seis módulos de trabajo distribuidos de la siguiente forma: (1) Introducción del programa, psicoeducación sobre la ansiedad, identificación de sintomatología y explorar la motivación al tratamiento (2) Intervención en las cogniciones relaciones con la ansiedad a través de la identificación de pensamientos automáticos (3) Entrenamiento en mindfulness y ejercicios de relajación (4) Comprender y practicar exposiciones (5) Intervención en habilidades sociales y relaciones interpersonales (6) Resumen y prevención de recaídas.

Velibra realiza la intervención a través de breves pasajes de texto, ilustraciones, grabaciones de audio, documentos PDF y mensajes de texto diarios recibidos como SMS o correo electrónico. A su vez posee cuestionarios que permiten llevar a cabo un seguimiento del síntoma, además ofrece la posibilidad de mantener un “diálogo individualizado” entre el programa y el usuario.

En el caso de Richards et al. (2015) habla de un programa de intervención en depresión distinto a lo expuesto con anterioridad. Expone el programa iCBT (terapia cognitivo-conductual informatizada) el cual consta de siete módulos en los que se trabajan tanto aspectos cognitivos como comportamentales. Los módulos a trabajar fueron los siguientes (1) Introducción al programa y psicoeducación sobre la depresión (2) Seguimiento del estado de ánimo y psicoeducación emocional (3) Cogniciones (4) Conductas. Activación conductual (5) Cuestionar pensamientos automáticos (6) Intervención en las creencias profundas (7) Módulo final de recogida de las habilidades e ideas adquiridas.

Además de los programas ya presentados, podemos mencionar Woebot (Fitzpatrick et al., 2017) una aplicación de mensajería instantánea para conversaciones diarias breves y seguimiento del estado de ánimo. Al inicio del programa se recopila información general sobre el estado de ánimo del usuario y se presentan los conceptos básicos relacionados con la terapia cognitivo-

conductual. Su software ha sido entrenado para generar respuestas empáticas atendiendo al estado anímico del participante, adaptar los contenidos a las características individualizadas, fijar los objetivos a alcanzar, trabajar la motivación y el compromiso del paciente. Además de ofrecer reflexiones semanales como forma de mostrar el estado anímico del participante a lo largo del tiempo.

En el caso de Tessa (Fitzsimmons-Craft et al., 2021) consiste en una primera fase introductoria además de ocho sesiones en las que se trabaja el ideal corporal relacionado con la delgadez, los medios de comunicación, las 4C (comparaciones, conversaciones, anuncios y ropa), alimentación saludable, comentarios críticos, ejercicio, atracones y mantenimiento. Además de los módulos predeterminados cuenta con un módulo “crisis” el cual se activa al considerar que el paciente puede encontrarse en riesgo, tratando de ofrecer asistencia ante la situación de dificultad.

Asimismo, Sandoval et al. (2017) hace uso de imBPST donde se trata de superar un punto débil observado en otras intervenciones cognitivo-conductual. Muestra un especial cuidado en desarrollar y mantener la alianza terapéutica a lo largo del tiempo (Barazzone et al., 2012). Esto lo hace ofreciendo un terapeuta virtual en pantalla con el que poder interactuar durante el uso del programa. Además, si el usuario muestra dificultades, el programa lanza preguntas específicas sobre el problema y proporciona una asistencia flexible para su resolución.

En último lugar Sadeh et al. (2023) presentan un formato de intervención distinto a lo mencionado con anterioridad, Eleos health platform es un programa que transcribe la información de las sesiones y manda un resumen al paciente. Este artículo se ha considerado de interés ya que puede acercar posturas a la hora de analizar si el uso de inteligencia artificial en la actualidad puede quitar o acompañar al trabajo llevado a cabo por un psicoterapeuta. Ofreciendo así una visión distinta sobre lo que es capaz de realizar la inteligencia artificial hasta el momento.

Impacto de las intervenciones

Analizando el impacto del uso de *chatbots* sobre los participantes, podemos observar como de forma general se obtiene una reducción en la sintomatología depresiva y/o ansiosa tras la intervención (**Tabla 3**).

En los resultados en síntomas depresivos, el programa Deprexis generalmente muestra un tamaño del efecto alto o moderado. De manera que aquellos participantes que hacen uso de este tipo de intervenciones, tiende a observar una disminución en la sintomatología depresiva.

Además del Deprexis, existen intervenciones como Amiria, centrada en esclerosis múltiple y sus síntomas (Gold et al., 2023) y Woebot, intervención con enfoque en las habilidades del terapeuta (Fitzpatrick et al., 2017) que no registran mejoras en los síntomas. Dicho de otra manera, el tamaño del efecto observado es negativo, lo cual muestra que tras las intervenciones no aparece una reducción en la sintomatología del participante.

Mencionando el uso de Woebot, es un programa centrado en el estado de ánimo y habilidades del terapeuta como la empatía y motivación del paciente. Y es que, a pesar de no mostrar cambios en síntomas depresivos, sí que muestra un tamaño del efecto moderado ($d = 0.56$) respecto a la disminución de síntomas ansiosos en los participantes.

En la línea de recogida de datos sobre síntomas ansiosos, en términos generales, se observa un tamaño del efecto moderado o pequeño. En el caso de Tessa – studentBodies (Fitzsimmons-Craft et al., 2021) y Mobile app-based interactive CBT (Oh et al., 2020), podemos observar resultados de efecto bajo ante ambas de las sintomatologías analizadas. Aunque no podemos alcanzar un fin concluyente, ya que la recogida de datos sobre este tipo de programas de intervención tan solo se lleva a cabo en un estudio. Para poder alcanzar conclusiones considerables se ha de contar con mayor número de estudios.

En términos generales y atendiendo a los resultados mostrados en la totalidad de estudios seleccionados para este metaanálisis, se podría considerar que los datos aportados en los estudios

sobre la reducción de síntomas depresivos es más consistente que lo observado sobre ansiedad. Ante estas conclusiones, puede estar incidiendo la diferencia en la proporción de estudios seleccionados de cada uno de los temas.

Con motivo de observar si estos cambios en los síntomas han sido mantenidos en el tiempo se ha llevado a cabo un análisis entre los resultados obtenidos previos a la intervención y tras seis meses de finalizarla. Cada estudio lleva a cabo una medida de seguimiento diferente, pudiendo ser a los dos meses (Sandoval et al., 2017), tres meses (Gold et al., 2023) e incluso a los tres y seis meses (Fitzsimmons-craft et al., 2021; Richards et al., 2015). Aunque para el análisis de este estudio y atendiendo a la medida predominante, tan solo se va a llevar a cabo un análisis respecto a los datos recogidos pasados los seis meses desde la intervención (Berger et al., 2011; Berger et al., 2017; Berger et al., 2018; Meyer et al., 2009; Schroder et al., 2014).

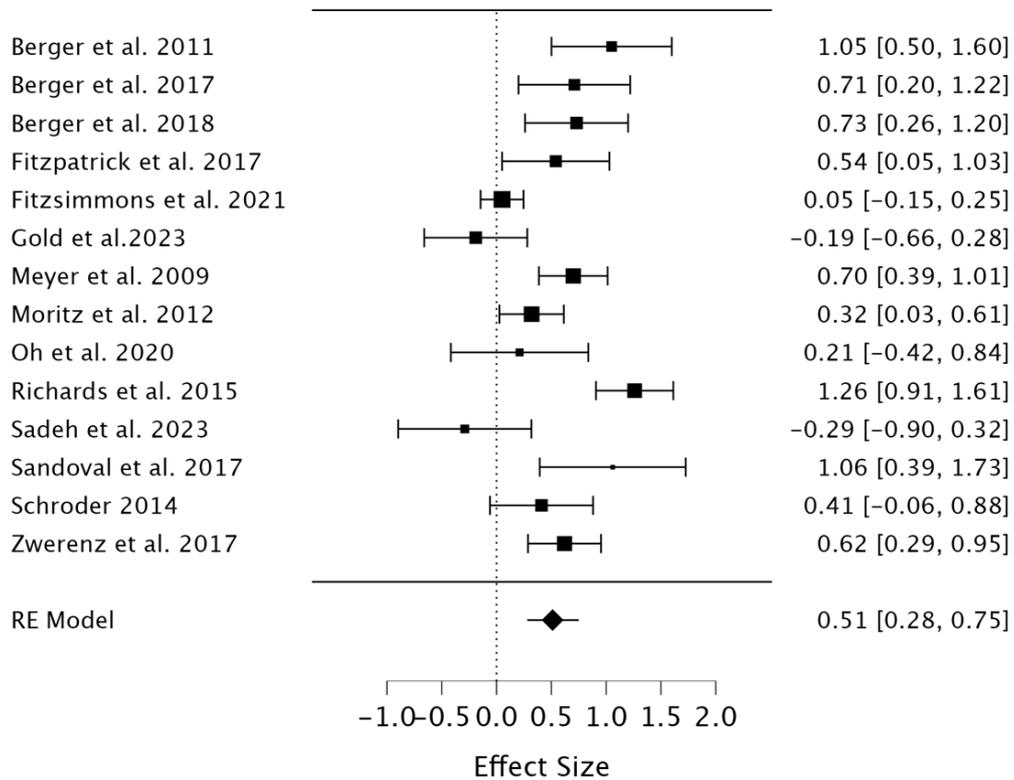
Por último, los resultados obtenidos en base a las características del grupo control (tratamiento habitual o lista de espera) no muestran grandes diferencias entre ellos. Bien es cierto que existe un mayor número de estudios en lista de espera o sin tratamiento que concluyen una reducción grande en sus síntomas, pero los datos recogidos en este estudio no se consideran suficientes como para poder alcanzar conclusiones determinantes al respecto.

Resultado del metaanálisis sobre la depresión

En los 14 estudios seleccionados para evaluar si la intervención por medio de *chatbots* es efectiva para reducir la sintomatología depresiva, se ha observado un tamaño de efecto medio de $d = 0.51$; $p < 0.001$; $IC95\% = [0.28 - 0.75]$ (**Figura 2**). Lo que nos indica que el grupo que recibe la intervención por medio *chatbots* reduce más la sintomatología depresiva que el grupo control. Encontramos una alta heterogeneidad entre los estudios ($Q_w = 64.1$ $p < 0.001$; $I^2 = 78.9\%$). Por lo que, hemos analizado el posible papel de distintos moderadores (**Tabla 2**).

Figura 2

Gráfico de metaanálisis entre la relación de uso de chatbots y reducción de síntomas depresivos

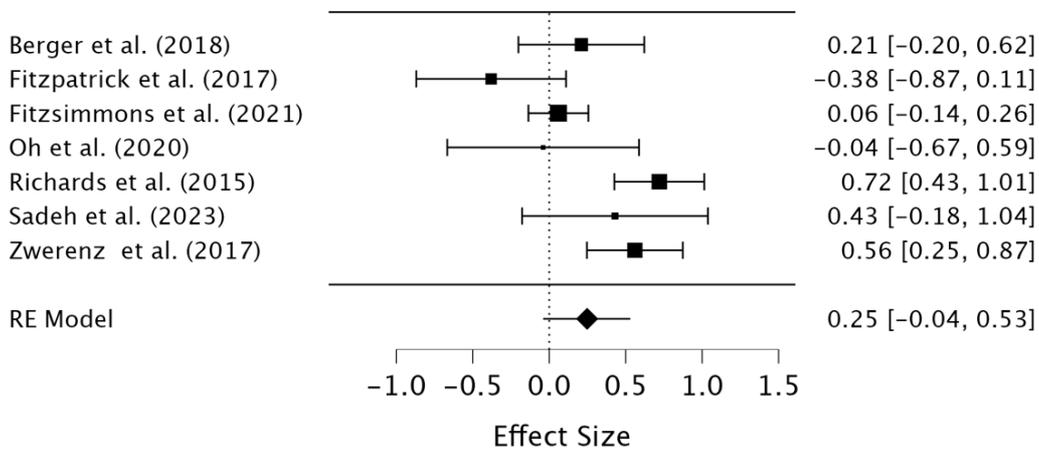


Resultado del metaanálisis sobre la ansiedad

En el metaanálisis clásico realizado con 7 estudios seleccionados, se observa un tamaño del efecto entre el uso de *chatbots* y la disminución en síntomas ansiosos de $d = 0.25$; $p = 0.09$; $IC95\% = [-0.04 - 0.53]$ (**Figura 3**). Los resultados alcanzados no permiten concluir que el uso de *bots* genere una reducción en los síntomas ansiosos del paciente. Entre estos estudios, existe una alta heterogeneidad ($Q_w = 24.7$, $p < 0.001$; $I^2 = 75.6\%$).

Figura 3.

Gráfico de metaanálisis entre la relación de uso de chatbots y reducción de sintomatología ansiosa



Análisis de moderadores

Se ha llevado a cabo un análisis de variables que moderan los resultados en la interacción entre programas informáticos y sintomatología depresiva ($Q_w = 64.1$; $p < 0.001$) y, con síntomas ansiosos ($Q_w = 24.7$; $p < 0.001$). Para ambas sintomatologías se ha analizado la edad media, género, población clínica o no clínica, el programa de intervención y el test de medida utilizado. Además, teniendo en cuenta uno de los objetivos del estudio, se ha analizado el tipo de grupo control entre los que se encuentra lista de espera (LDE) o tratamiento habitual (TH).

En el análisis de moderadores en síntomas depresivos, generalmente se han recogido factores no significativos, como la edad media ($Q_b = 0.3$; $p = 0.58$), género ($Q_b = 2.39$; $p = 0.12$), semanas de intervención ($Q_b = 0.1$; $p = 0.73$), población clínica ($Q_b = 2.86$; $p = 0.09$), programa de intervención ($Q_b = 0.7$; $p = 0.40$) y tipo de grupo control ($Q_b = 1.53$; $p = 0.22$). Entre todas estas variables, el tipo de test utilizado en la medida muestra resultados estadísticamente significativos ($Q_b = 4.14$; $p = 0.04$).

Todas las variables moderadoras analizadas en síntomas ansiosos destacaron ser no significativos, edad media ($Q_b = 3.64$; $p = 0.06$), género ($Q_b = 0.08$; $p = 0.78$), semanas de intervención ($Q_b = 3.32$; $p = 0.07$), población clínica ($Q_b = 0.0$; $p = 0.96$), programa de

intervención ($Qb = 0.42; p = 0.51$), tipo de grupo control ($Qb = 0.37; p = 0.54$) y test de medida ($Qb = 0.43; p = 0.51$).

Tabla 2

Análisis de moderadores de en síntomas depresivos y ansiosos

	DEPRESIÓN	ANSIEDAD
	$Q_w = 64.1 p < 0.001$	$Q_w = 24.7 p < 0.001$
<i>Edad media</i>	$Q_b = 0.3; p = 0.58$	$Q_b = 3.64; p = 0.06$
<i>Género</i>	$Q_b = 2.39; p = 0.12$	$Q_b = 0.08; p = 0.78$
<i>Semanas de intervención</i>	$Q_b = 0.1; p = 0.73$	$Q_b = 3.32; p = 0.07$
<i>Población clínica</i>	$Q_b = 2.86; p = 0.09$	$Q_b = 0.0; p = 0.96$
<i>Programa de intervención</i>	$Q_b = 0.7; p = 0.40$	$Q_b = 0.42; p = 0.51$
<i>Tipo de grupo control</i>	$Q_b = 1.53; p = 0.22$	$Q_b = 0.37; p = 0.54$
<i>Test de medida</i>	$Q_b = 4.14; p = 0.04$	$Q_b = 0.43; p = 0.51$

Tabla 3*Resultados de las intervenciones*

Autor (año)	Nº de ramas	Programa	Comparación	Medida	Pre-post*	Seguimiento**
Berger et al. (2011)	3	Deprexis	LDE	BDI-II	1.07	0.37
Berger et al. (2017)	2	Velibra iCBT	TH IP	BDI-II	0.72	
Berger et al. (2018)	2	Deprexis	TH IP	BDI-II GAD-7	0.74 0.21 ^a	0.64 0.15 ^a
Fitzpatrick et al. (2017)	2	Woebot	TH Lecturas	PHQ-9 GAD-7	-0.39 0.56 ^a	
Fitzsimmons-Craft et al. (2021)	3	Tessa StudentBodies	LDE	PHQ-8 GAD-7	0.05 0.07 ^a	0.04 0.05 ^a
Gold et al. (2023)	3	Amiria iCBT	TH IP	BDI-II	-0.20	
Meyer et al. (2009)	2	Deprexis	LDE	BDI-II	0.70	-0.28
Moritz et al. (2012)	2	Deprexis	LDE	BDI-II	0.32	
Oh et al. (2020)	2	Mobile app-based interactive CBT	TH Lecturas	HADS HADS	0.22 -0.05 ^a	
Richards et al. (2015)	2	iCBT	LDE	BDI-II GAD-7	1.27 0.72 ^a	
Sadeh et al. (2023)	2	Eleos health platform	TH IP	PHQ-9 GAD-7	0.27 0.44 ^a	
Sandoval et al. (2017)	2	imbPST	NT	BDI-II	1.10	
Schroder et al. (2014)	2	Deprexis	LDE	BDI-II	0.42	
Zwerenz et al. (2017)	2	Deprexis	TH ISO	BDI-II PHQ-9 GAD-7	0.63 0.56 0.40 ^a	

* = Tamaño del efecto (*d* de Cohen) – Interacción pre-post

** = Tamaño del efecto (*d* de Cohen) – Interacción pre-seguimiento

^a = ansiedad

LDE = Lista de espera; TH= Tratamiento habitual; NT = No tratamiento; IP = Intervención personal (cara a cara); ISO= Información semanal online

DISCUSIÓN

En los últimos años, el uso de las nuevas tecnologías ha adquirido un papel fundamental en el día a día de la población. Esta tendencia se ha visto derivada no solo por la constante evolución tecnológica, sino también por la crisis sanitaria generada por la COVID-19. Durante este periodo muchas personas experimentaron un gran aislamiento al haber sido confinados y distanciados de su entorno social habitual. Todo ello resultó en un aumento notable en casos de depresión y ansiedad, además de dificultades en las relaciones sociales (Brooks et al., 2020).

Como forma de hacer frente al malestar generado por la pandemia, la inteligencia artificial mostró ser una herramienta valiosa. Su uso ha permitido generar nuevas formas de interacción social dando como resultado un mayor sentimiento de conexión con las personas. Aunque también resultó en estilos de vida más sedentarios a causa del incremento de actividades desde casa como el teletrabajo y el entretenimiento digital. Esto ha generado cambios en el ámbito social, personal y laboral de las personas, influyendo en su jornada de trabajo, las interacciones sociales y en el bienestar tanto físico como mental (Wiederhold, 2020).

A lo largo de este metaanálisis se ha tratado de analizar el impacto de programas de inteligencia, en específico de plataformas *chatbots* como forma de hacer frente al aumento de sintomatología depresiva y/o ansiosa. Donde los *bots* son definidos como programas informáticos diseñados para llevar a cabo conversaciones inteligentes y naturales con seres humanos (Pérez y Céspedes, 2021). Además, se ha analizado el mantenimiento de los resultados a lo largo del tiempo y posibles moderadores que pudiesen estar incidiendo en los resultados alcanzados.

En cuanto a los resultados de este estudio, se ha observado una correlación positiva entre el uso de *bots* y la reducción en síntomas depresivos. Se ha observado una disminución

en síntomas depresivos tras recibir una intervención online a través de plataformas *chatbot*. Aunque, estos mismos resultados no se sostienen en el análisis de síntomas ansiosos, donde no se concluye una reducción en los síntomas tras su uso. En ambas sintomatologías se han analizado variables moderadoras que pudiesen estar incidiendo sobre los resultados alcanzados. Entre ellas se destacan variables de la muestra como la edad, sexo, población clínica o no clínica. Y además, variables metodológicas como las semanas y programas de intervención, test de medida utilizado y tipo de grupo control, pudiendo ser aquellos sin ningún tipo de intervención (lista de espera) o con su tratamiento psicoterapéutico habitual (TH).

Entre las variables estudiadas no se han encontrado resultados significativos en ninguna de ellas. Exceptuando el tipo de test utilizado, donde se concluye que el test de medida afecta en los resultados de mejora de síntomas depresivos del paciente.

Analizando en profundidad los resultados en síntomas depresivos, autores como Burns et al. (2011) y Watts et al. (2013) en sus estudios también concluyen una disminución en signos y síntomas depresivos tras 8 semanas de participación en programas online. A su vez, metaanálisis previos como Twomey et al. (2017) y Wahle et al. (2017) alcanzaron las mismas conclusiones que en este análisis. El número de estudios que muestra un cambio en el tamaño del efecto podría llevar a plantear la consistencia de las intervenciones a través de *chatbots*.

Respecto a las variables estudiadas en depresión, el tipo de test utilizado mostraba resultados significativos. Se ha de tener en cuenta que generalmente las aplicaciones informáticas han sido diseñadas desde un enfoque cognitivo-conductual, mostrando una alta directividad y estructura orientada a los objetivos (Nordgreen et al., 2010, p.173). Tienden a ser programados para imitar conversaciones terapéuticas, entre los que se incluye la reestructuración cognitiva o la activación conductual como forma de mejorar las cogniciones negativas (D'Alfonso et al., 2017; Fernández y Mairal, 2017). Y es que teniendo en cuenta

que los participantes evaluados con el BDI-II han mostrado mayores resultados en el tamaño del efecto que aquellos evaluados con el PHQ-9, se ha de tomar en consideración sus similitudes y diferencias.

Revisando en detalle los ítems de ambos cuestionarios, se podría mencionar ciertas diferencias en ellos. El PHQ-9 es compuesto por nueve ítems en los que se evalúan síntomas fisiológicos del paciente que pudiesen estar asociados a rasgos depresivos (Kroenke, Spitzer, y Williams, 2001). En el caso del BDI-II, es un test desarrollado por Aaron T. Beck et al. (1996), autor conocido como uno de los padres de la terapia cognitiva. En sus líneas de trabajo sobre depresión, sugiere que los pensamientos automáticos negativos y las distorsiones cognitivas juegan un papel crucial en el desarrollo y mantenimiento de estos cuadros sintomáticos. El BDI-II se compone por un total de 21 ítems con una mirada más centrada en cuestiones cognitivas. Todo ello pudiese resultar en el BDI como un cuestionario sensible en la detección de cambios, especialmente en aquellos generados tras una intervención centrada en cogniciones y conductas, como son los *chatbots*.

La sensibilidad del BDI ha sido evaluada en varias investigaciones, Beck, A. T., et al. (1961) en sus estudios mostraban una sensibilidad del 93% a la hora de detectar la depresión clínica en pacientes psiquiátricos. Posteriormente Wang, Y. y Gorenstein, C. (2013) realizaron un metaanálisis donde se examinó la sensibilidad del test en diferentes poblaciones, dando un resultado promedio del 79%, lo que sugiere una buena capacidad para identificar casos de depresión en general. De manera que basándonos en estos estudios llevados a cabo, podemos concluir que el Inventario de Depresión de Beck (BDI) tiene una sensibilidad generalmente alta para detectar la presencia de depresión en diversas poblaciones, convirtiéndolo así en una herramienta útil en la evaluación y diagnóstico de depresión.

Por otro lado, se ha tratado de dar respuesta al mantenimiento de los resultados en el tiempo, donde se han observado tamaños del efecto bajo. No se han alcanzado datos

concluyentes acerca del impacto del uso de *chatbots* en síntomas depresivos y/o ansiosos a largo plazo. Estos datos eran esperables debido a la falta de continuidad de participantes, pérdida de sujetos, y además, el bajo número de estudios que llevan a cabo seguimiento en el tiempo. Todo esto podría estar limitando los efectos a largo plazo. El autor Karyotaki et al. (2017) concluyó que las intervenciones para la depresión basadas en internet mostraban mejoras a corto plazo, pero los efectos a largo plazo mostraron ser menos consistentes. Esto podría estar sugiriendo el beneficio de llevar a cabo intervenciones más intensivas o personalizadas para mantener los beneficios a lo largo del tiempo.

Analizando el último de los objetivos de este estudio, se ha tratado de concluir sobre la diferencia en grupos control “lista de espera” y aquellos que recibían tratamiento psicoterapéutico habitual. En este estudio no se han encontrado resultados estadísticamente significativos para poder concluir aspectos diferenciales entre los grupos. Incluso muestran resultados distinto a lo esperado. Ya que, atendiendo las conclusiones alcanzadas por otros estudios como Berger et al. (2018) y Richards y Richardson (2012), el uso de programas de autoayuda online usados como una herramienta adicional a la psicoterapia cara a cara, resulta en mayor éxito terapéutico. Ya que deriva en un aumento en la frecuencia e intensidad de sesiones de psicoterapia (Abramowitz et al., 2003; Reese et al., 2011).

Teniendo en cuenta los resultados inesperados en el estudio de grupos control, para próximas líneas de investigación se sugiere llevar a cabo amplios estudios en el que el grupo experimental pueda recibir psicoterapia a través de *chatbots*, mientras que un grupo control no reciba ninguna intervención alternativa. Esto nos permitiría alcanzar resultados con mayor validez.

Además, se sugiere continuar con el estudio del impacto de las intervenciones a largo plazo. De manera que se puedan alcanzar conclusiones consistentes sobre el mantenimiento de los resultados terapéuticos tras hacer uso de programas de inteligencia artificial. Incluso

pudiendo llegar a comparar estos con una intervención psicoterapéutica convencional. Esto también podría dar lugar a examinar la aceptabilidad y satisfacción del paciente sobre el uso de *chatbots* en psicoterapia.

Como última propuesta, resulta de especial interés poder investigar sobre los distintos modelos de psicoterapia y su alcance a través de *chatbots*. Evaluando así si existen diferencias en los resultados dependiendo del enfoque de intervención. Esto permite evaluar si todos los enfoques mantienen la misma eficacia, tanto cara a cara como a través del uso de nuevas tecnologías.

El presente análisis posee limitaciones las cuales se han de tener en cuenta. En primer lugar y atendiendo al crecimiento reciente de las nuevas tecnologías, la disponibilidad de datos científicos en relación a la inteligencia artificial es limitada. Con motivo de hacer frente a esto, se han seleccionado 14 artículos los cuales comparten similitudes en la intervención, aunque también muestran variaciones entre ellos. Se ha de mencionar que seis de los seleccionados centraban su estudio exclusivamente en sintomatología depresiva, mientras siete abordaban tanto depresión como ansiedad. De manera que no se han analizado estudios únicamente centrados en la intervención en síntomas ansiosos.

En cuanto a las variables de la muestra, se han encontrado disparidades. Entre los seleccionados podemos encontrarnos tanto con participantes clínicos como no clínicos, lo cual muestra una amplia variedad en la gravedad de los síntomas depresivos y/o ansiosos. Esto podría representar un impacto significativo en los resultados del estudio. Además de las diferencias en el tamaño de la muestra, abarcando desde 41 a 700 participantes.

Finalmente, se ha de mencionar las variables de intervención, donde la duración de las intervenciones mostraban ser entre 2 y 12 semanas. O en cuanto al seguimiento, se analizan estudios de no seguimiento, seguimiento a los 3 meses y/o tras 6 meses. La heterogeneidad en las intervenciones dificulta la comparación de resultados entre estudios. Además, se identificó

un alta tasa de abandono del tratamiento, lo cual puede sesgar los resultados y afectar de manera directa en la validez de las conclusiones obtenidas.

A pesar de las limitaciones indicadas anteriormente, los resultados muestran la importancia del tema estudiado. Por lo tanto, se considera relevante continuar con la investigación en este ámbito. El uso de *chatbots* se presenta como forma eficaz de paliar el malestar en la población. Su uso puede reducir el malestar y, a su vez, los costes económicos, pero es fundamental que no se logre restando importancia al sufrimiento de las personas.

Conclusiones

- Los programas informáticos destacan ser avances prometedores en el mundo de la salud mental.
- El uso de programas de intervención en específico *chatbots* entrenados desde un modelos cognitivo-conductual, tiende a reducir la sintomatología depresiva.
- La intervención a través de *chatbots* no tiende a mostrar gran impacto en la reducción de síntomas ansiosos.
- Se muestran mejores resultados en la reducción de síntomas depresivos al hacer uso del programa informático *Deprexis*.
- En la intervención online, no podemos afirmar que exista un mantenimiento de los resultados de mejora en el tiempo.

Referencias

- Abramowitz, J. S., Foa, E. B., & Franklin, M. E. (2003). Exposure and ritual prevention for obsessive-compulsive disorder: Effects of intensive versus twice-weekly sessions. *Journal of Consulting and Clinical Psychology, 71*(2), 394–398.
- Addis, M. E., Truax, P., & Jacobson, N. S. (1995). Why do people think they are depressed? The Reasons For Depression Questionnaire. *Psychotherapy: Theory, Research, Practice, Training, 32*(3), 476-483.
- Ahuja, A. S. (2019). The impact of artificial intelligence in medicine on the future role of the physician: A study [Doctoral dissertation, Name of Institution].
- Amann, J., Blasimme, A., Vayena, E., Frey, D., Madai, V. I., & Precise4Q Consortium. (2020). Explainability for artificial intelligence in healthcare: A multidisciplinary perspective. *BMC Medical Informatics and Decision Making, 20*, 1-9. <https://doi.org/10.1186/s12911-020-1089-7>
- Aparicio Gómez, W. O. (2023). La Inteligencia Artificial y su Incidencia en la Educación: Transformando el Aprendizaje para el Siglo XXI. *Revista Internacional De Pedagogía E Innovación Educativa, 3*(2), 217–229. <https://doi.org/10.51660/ripie.v3i2.133>
- Aretio, L. G. (2021). COVID-19 y educación a distancia digital: Preconfinamiento, confinamiento y posconfinamiento. *RIED. Revista Iberoamericana de Educación a Distancia, 24*(1), 9-32.
- Astobiza, A. M. (2020). Medicina digital y el futuro de la salud. *Dilemata, (32)*, 5-16.
- Ausina, J. B., & Meca, J. S. (2015). *Meta-análisis en ciencias sociales y de la salud*. Editorial Síntesis.
- Ávila-Tomás, J. F., Mayer-Pujadas, M. A., & Quesada-Varela, V. J. (2020). La inteligencia artificial y sus aplicaciones en medicina I: Introducción antecedentes a la IA y robótica. *Atención Primaria, 52*(10), 778–784. <https://doi.org/10.1016/j.aprim.2020.04.013>
- Balsategui, A. R. (2019). La subjetividad del terapeuta. *Norte de Salud Mental, 16*(60), 21-32.
- Barazzone, N., Cavanagh, K., & Richards, D. A. (2012). Computerized cognitive behavioural therapy and the therapeutic alliance: A qualitative enquiry. *British Journal of Clinical Psychology, 51*(4), 396–417.

- Baumeister, H., Reichler, L., Munzinger, M., & Lin, J. (2017). The impact of guidance on Internet-based mental health interventions — A systematic review. *Internet Interventions*, 8, 1–10. <https://doi.org/10.1016/j.invent.2017.04.004>
- Beck, A. T. (1970). Cognitive therapy: Nature and relation to behavior therapy. *Behavior Therapy*, 1(2), 184–200. [https://doi.org/10.1016/S0005-7894\(70\)80030-2](https://doi.org/10.1016/S0005-7894(70)80030-2)
- Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Manual for the Beck Depression Inventory-II*. San Antonio, TX: Psychological Corporation.
- Beck, A. T., Ward, C. H., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry*, 4(6), 561–571. <https://doi.org/10.1001/archpsyc.1961.01710120031004>
- Benbya, H., Davenport, T. H., & Pachidi, S. (2020). Artificial intelligence in organizations: Current state and future opportunities. *MIS Quarterly Executive*, 19(4), 301-316.
- Benítez, R., Escudero, G., Kanaan, S., & Massip, R. (2014). *Inteligencia Artificial Avanzada*. Editorial UOC.
- Botella Ausina, J., & Caperos Montalbán, J. M. (2019). *Metodología de Investigación en Psicología General Sanitaria. Ciencias sociales y humanidades* (Vol. 9). Editorial.
- Botella, J., & Sánchez-Meca, J. (2015). *Meta-análisis en ciencias sociales y de la salud*. Editorial.
- Brooks, S. K., Webster, R. K., Smith, L. E., Woodland, L., Wessely, S., Greenberg, N., & Rubin, G. J. (2020). The psychological impact of quarantine and how to reduce it: rapid review of the evidence. *The Lancet*, 395(10227), 912-920.
- Burns, M. N., Begale, M., Duffecy, J., Gergle, D., Karr, C. J., Giangrande, E., & Mohr, D. C. (2011). Harnessing context sensing to develop a mobile intervention for depression. *Journal of Medical Internet Research*, 13(3), e55. <https://doi.org/10.2196/jmir.1838>
- Castellaro, M. (2011). El concepto de representación mental como fundamento epistemológico de la psicología. *Límite: Revista de Filosofía y Psicología*, (24), 55-68.
- Cedeño, N. J. V., Cuenca, M. F. V., Mojica, Á. A. D., & Portillo, M. T. (2020). Afrontamiento del COVID-19: Estrés, miedo, ansiedad y depresión. *Enfermería Investiga*, 5(3), 63-70.
- Chavez Bazan, F. L. (2020). *Inteligencia artificial en tratamientos fóbicos: Una revisión de la literatura científica (2010–2020)*.

- Clark, L., & Watson, D. (1995). Testing a tripartite model: Exploring the symptom structure of anxiety and depression in student, adult, and patient samples. *Journal of Abnormal Psychology, 104*(1), 15-25.
- Cohen, J. (2013). *Statistical Power Analysis for the Behavioral Sciences* (0 ed.). Routledge. <https://doi.org/10.4324/9780203771587>
- Confederación de la Salud Mental en España & Fundación Mutua Madrileña. (2023). La situación de la salud mental en España. Recuperado de <https://www.consaludmental.org/publicaciones/Estudio-situacion-salud-mental-2023.pdf>
- Coyne, J. C. (1976). Toward an interactional description of depression. *Psychiatry, 39*(1), 28-40.
- Cuijpers, P., van Straten, A., & Warmerdam, L. (2007). Behavioral activation treatments of depression: A meta-analysis. *Clinical Psychology Review, 27*(3), 318-326. <https://doi.org/10.1016/j.cpr.2006.11.001>
- D'Alfonso, S., Santesteban-Echarri, O., Rice, S., Wadley, G., Lederman, R., Miles, C., Gleeson, J., & Alvarez-Jimenez, M. (2017). Artificial intelligence-assisted online social therapy for youth mental health. *Frontiers in Psychology, 8*, 796. <https://doi.org/10.3389/fpsyg.2017.00796>
- D'Zurilla, T. J., & Goldfried, M. R. (1971). Problem solving and behavior modification. *Journal of Abnormal Psychology, 78*(1), 107–126. <https://doi.org/10.1037/h0031360>
- Dunn, A. L., Trivedi, M. H., Kampert, J. B., Clark, C. G., & Chambliss, H. O. (2005). Exercise treatment for depression: Efficacy and dose response. *American Journal of Preventive Medicine, 28*(1), 1-8. <https://doi.org/10.1016/j.amepre.2004.09.003>
- Duro Martínez, J. C. (2021). ¿Sabemos cuántos profesionales especialistas en psicología clínica trabajan en el Sistema Nacional de Salud español?. *Papeles del Psicólogo, 42*(2), 81-93.
- Fernández, E. N., & Mairal, J. B. (2017). Behavioral activation versus cognitive restructuring to reduce automatic negative thoughts in anxiety generating situations. *Psicothema, 29*(2), 172–177.
- Fiske, A., Henningsen, P., & Buyx, A. (2019). Your robot therapist will see you now: Ethical implications of embodied artificial intelligence in psychiatry, psychology, and

- psychotherapy. *Journal of Medical Internet Research*, 21(5), e13216. <https://doi.org/10.2196/13216>
- Fitzpatrick, K. K., Darcy, A., & Vierhile, M. (2017). Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (Woebot): A randomized controlled trial. *JMIR Mental Health*, 4(2), e19. <https://doi.org/10.2196/mental.7785>
- Friedler, S. A., Scheidegger, C., & Venkatasubramanian, S. (2019). The challenges of machine 'error' and 'bias': Diverse perspectives from the machine learning community. *Big Data*, 7(2), 92-101. <https://doi.org/10.1089/big.2019.29030.saf>
- Gable, S. L., Reis, H. T., Impett, E. A., y Asher, E. R. (2004). What do you do when things go right? The intrapersonal and interpersonal benefits of sharing positive events. *Journal of Personality and Social Psychology*, 87(2), 228-245. <https://doi.org/10.1037/0022-3514.87.2.228>
- George, A. S., & George, A. H. (2023). A review of ChatGPT AI's impact on several business sectors. *Partners Universal International Innovation Journal*, 1(1), 9-23.
- Gobierno de España. (2023, 5 de junio). Conoce la Estrategia Nacional de Inteligencia Artificial (ENIA). Recuperado de <https://planderecuperacion.gob.es/noticias/conoce-Estrategia-Nacional-Inteligencia-Artificial-ENIA-IA-prtr>
- Gutiérrez Martínez, J. A., & Febles Estrada, A. (2020). Las tecnologías disruptivas y su aplicación en la medicina con vistas al 2030. *Revista Cubana de Salud Pública*, 45, e1563.
- Harkness, K. L., & Monroe, S. M. (2002). Childhood adversity and the endogenous versus nonendogenous distinction in women with major depression. *American Journal of Psychiatry*, 159(3), 387-393. <https://doi.org/10.1176/appi.ajp.159.3.387>
- Hautzinger, M. (2003). *Kognitive Verhaltenstherapie bei Depressionen (Cognitive behavioural therapy of depression)*. Verlagsgruppe Beltz.
- Hayes, S. C., Strosahl, K., & Wilson, K. G. (1999). *Acceptance and Commitment Therapy: An Experiential Approach to Behavior Change*. Guilford Press.
- Higgins, J. P. T., y Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*, 21(11), 1539-1558. <https://doi.org/10.1002/sim.1186>

- Hill, C. E., Crook-Lyon, R. E., & Hess, S. A. (2006). Prediction of session process and outcome in the Hill dream model: Contributions of client characteristics and the process of the three stages. *Dreaming*, 16(3), 159-185. <https://doi.org/10.1037/1053-0797.16.3.159>
- Hofmann, S. G., Asnaani, A., Vonk, I. J., Sawyer, A. T., & Fang, A. (2012). The efficacy of cognitive behavioral therapy: A review of meta-analyses. *Cognitive Therapy and Research*, 36(5), 427–440. <https://doi.org/10.1007/s10608-012-9476-1>.
- Hollon, S. D., Thase, M. E., & Markowitz, J. C. (2002). Treatment and prevention of depression. *Psychological Science in the Public Interest*, 3(2), 39-77. <https://doi.org/10.1111/1529-1006.00008>
- Hood, L., & Friend, S. H. (2011). Predictive, personalized, preventive, participatory (P4) cancer medicine. *Nature Reviews. Clinical Oncology*, 8(3), 184–187. <https://doi.org/10.1038/nrclinonc.2010.227>
- Huedo-Medina, T. B., Sánchez-Meca, J., Marín-Martínez, F., y Botella, J. (2006). Assessing heterogeneity in meta-analysis: Q statistic or I2 index? *Psychological Methods*, 11(2), 193-206. <https://doi.org/10.1037/1082-989X.11.2.193>
- JASP Team. (Año). JASP (Versión 0.18.3) [Software de computadora]. Recuperado de <https://jasp-stats.org/>
- Joiner, T. E. (2000). Depression's vicious scree: self-propagating and erosive processes in depression chronicity. *Clinical Psychology: Science and Practice*, 7, 203-218. <https://doi.org/10.1093/clipsy.7.2.20>
- Kroenke, K., Spitzer, R. L., & Williams, J. B. (2001). The PHQ-9: Validity of a brief depression severity measure. *Journal of General Internal Medicine*, 16(9), 606-613. <https://doi.org/10.1046/j.1525-1497.2001.016009606.x>
- Lambert, M. (1992). Implications of psychotherapy outcome research for eclectic and integrative psychotherapies. En S. L. Garfield y A. E. Bergin (Eds.), *Handbook of psychotherapy and behavior change* (pp. 157-212). Wiley.
- Lambert, M. J., Shapiro, D. A., & Bergin, A. E. (1986). The effectiveness of psychotherapy. En S. L. Garfield y A. E. Bergin (Eds.), *Handbook of psychotherapy and behavior change* (pp. 157-212). Wiley.

- Lambert, S., McCreary, B., Joiner, T., Schmidt, N., & Ialongo, N. (2004). Structure of Anxiety and Depression an Examination of the Tripartite Model. *Journal of Consulting and Clinical Psychology*, 72(5), 904-908.
- Lejuez, C. W., Hopko, D. R., & Hopko, S. D. (2001). A brief behavioral activation treatment for depression: Treatment manual. *Behavior Modification*, 25(2), 255-286. <https://doi.org/10.1177/0145445501252005>
- Lin, J., Paganini, S., Sander, L., Lu'king, M., Ebert, D. D., Buhrman, M., Andersson, G., & Baumeister, H. (2017). An Internet-based intervention for chronic pain: A three- arm randomized controlled study of the effectiveness of guided and unguided acceptance and commitment therapy. *Deutsches Ärzteblatt International*, 114(41), 681.
- Lindhiem, O., Bennett, C. B., Rosen, D., & Silk, J. (2015). Mobile technology boosts the effectiveness of psychotherapy and behavioral interventions: A meta-analysis. *Behavior Modification*, 39(6), 785–804. <https://doi.org/10.1177/0145445515595198>.
- Mahoney, J. (s.f.). La teoría del procesamiento de la información de Mahoney. RODAS. Repositorio de Objetos de Aprendizaje de la Universidad de Sevilla. Recuperado el 4 de marzo de 2019, de https://rodas5.us.es/file/1240b064-8389-6228-96a5-653dd137f73b/1/capitulo3_SCORM.zip/pagina22.htm
- Ministerio de Sanidad. (2020). Encuesta Europea de Salud en España 2020: Informe de resultados. https://www.sanidad.gob.es/estadEstudios/estadisticas/EncuestaEuropea/EncuestaEuropea2020/EESE2020_inf_evol_princip_result.pdf
- Mitchell, T. (2006). The discipline of machine learning. Machine Learning Department technical report CMU-ML-06-108.
- Morris, J. (1985). *The Dream Workbook: Discover the Knowledge and Power Hidden in your Dreams*. Fawcett Crest.
- Motlagh, N. Y., Khajavi, M., Sharifi, A., & Ahmadi, M. (2023). The impact of artificial intelligence on the evolution of digital education: A comparative study of openAI text generation tools including ChatGPT, Bing Chat, Bard, and Ernie. arXiv preprint arXiv:2309.02029.
- Nordgreen, T., Havik, O. E., Ost, L. G., Furmark, T., Carlbring, P., Andersson, G., & Heiervang, E. (2010). Evaluating the therapeutic relationship in a cognitive-behavioral

- e-therapy program for generalized anxiety disorder: Development and pilot testing. *Journal of Medical Internet Research*, 12(5), e56. <https://doi.org/10.2196/jmir.1481>
- OMS. (2022, 2 de marzo). Salud mental y COVID-19: Datos iniciales sobre las repercusiones de la pandemia. https://www.who.int/es/publications/i/item/WHO-2019-nCoV-Sci_Brief-Mental_health-2022.1
- Patel, V., Maj, M., Flisher, A. J., De Silva, M. J., Koschorke, M., Prince, M., Zonal, W., Representatives, M. S., Tem- pier, R., Riba, M., & Sanchez, M. (2010). Reducing the treatment gap for mental disorders: A WPA survey. *World Psychiatry*, 9(3), 169–176. <https://doi.org/10.1002/j.2051-5545.2010.tb00305.x>.
- Paz-y-Miño, C. (2023). La inteligencia artificial en medicina general y en genómica. *Metro Ciencia*, 31(2), 81–86. <https://doi.org/10.47464/MetroCiencia/vol31/2/2023/81-86>
- Pérez, M. H., & Céspedes, L. Á. L. (2021). Definición de un proceso ingenieril para el desarrollo de un chatbot a partir de buenas prácticas establecidas. *Revista Cubana de Transformación Digital*, 2(3), 90-109.
- Reese, R. J., Toland, M. D., & Hopkins, N. B. (2011). Replicating and extending the good-enough level model of change: Considering session frequency. *Psychotherapy Research*, 21, 608–619.
- Richards, D., & Richardson, T. (2012). Computer-based psychological treatments for depression: A systematic review and meta-analysis. *Clinical Psychology Review*, 32(4), 329-342. <https://doi.org/10.1016/j.cpr.2012.02.004>
- Richardson, C. R., Avripas, S. A., Neal, D. L., & Marcus, S. M. (2005). Increasing lifestyle physical activity in patients with depression or other serious mental illness. *Journal of Psychiatric Practice*, 11(6), 379-388. <https://doi.org/10.1097/00131746-200511000-00004>
- Rivera, E. J. E., & Sánchez, S. D. V. (2016). Inteligencia artificial ¿reemplazando al humano en la psicoterapia? *Escritos*, 24(53), 271–291. <https://doi.org/10.18566/escr.v24n53.a02>
- Rodríguez, J. M., Merlino, H., & Fernández, E. (2014). Comportamiento adaptable de chatbots dependiente del contexto. *Archivo de la revista Latinoamericana de Ingeniería de Software*, 2(2), 115-136.

- Rohde, P., Lewinsohn, P., & Seeley, J. (1991). Comorbidity of unipolar depression with other mental disorders in adolescents and adults. *Journal of Abnormal Psychology*, 100(2), 214-222.
- Romero, M., Casadevante, C., & Montoro, H. (2020). Cómo construir un psicólogo-chatbot. *Papeles del Psicólogo*, 41(1), 27-34.
- Rudolph, J., Tan, S., & Tan, S. (2023). ChatGPT: Bullshit spewer or the end of traditional assessments in higher education?. *Journal of Applied Learning and Teaching*, 6(1).
- Ruiz-Rodríguez, P., Cano-Vindel, A., Muñoz Navarro, R., Medrano, L., Moriana, J. A., Buiza Aguado, C., Jiménez Cabré, G., & González-Blanch, C. (2017). Impacto económico y carga de los trastornos mentales comunes en España: una revisión sistemática y crítica. *Ansiedad y estrés*, 23(2-3), 118-123. <https://doi.org/10.1016/j.anyes.2017.10.003>
- Ruiz, J., & Cano, J. (2002). *Manual de psicoterapia cognitiva*. R&C Editores.
- Ruiz, R. B., & Velásquez, J. D. (2022). Inteligencia artificial al servicio de la salud del futuro. *Revista Médica Clínica las Condes*, 34(1), 84-91.
- Safran, J., & Segal, Z. (1994). *El proceso interpersonal en la terapia cognitiva*. Paidós.
- Schnyder, N., Panczak, R., Groth, N., & Schultze-Lutter, F. (2017). Association between mental health-related stigma and active help-seeking: Systematic review and meta-analysis. *British Journal of Psychiatry*, 210(4), 261-268. <https://doi.org/10.1192/bjp.bp.116.189464>.
- Segal, Z. V., Williams, J. M. G., & Teasdale, J. (2002). *Mindfulness-Based Cognitive Therapy for Depression: A New Approach to Preventing Relapse*. Guilford.
- Seidmann, S. (2000). *Historia de la psicología social*. https://www.psi.uba.ar/academica/carrerasdegrado/psicologia/sitios_catedras/obligatorias/035_psicologia_social1/material/descargas/historia_psico_social.pdf
- Seligman, M. E. P., Rashid, T., & Parks, A. C. (2006). Positive psychotherapy. *American Psychologist*, 61(8), 774-788. <https://doi.org/10.1037/0003-066X.61.8.774>
- Selma Penalva, A. (2021). Inteligencia artificial y Derecho del Trabajo. *Ius et Scientia*, 7(2), 29-40.

- Shalev-Shwartz, S., & Ben-David, S. (2014). *Understanding machine learning: From theory to algorithms*. Cambridge University Press.
- Suta, P., Lan, X., Wu, B., Mongkolnam, P., & Chan, J. H. (2020). An overview of machine learning in chatbots. *International Journal of Mechanical Engineering and Robotics Research*, 9(4), 502-510.
- Taylor, J. K., & Kowalczyk, J. (2020). Patient trust in artificial intelligence-enabled telemedicine in diabetes management. *The Journal of Medical Internet Research*, 22(2), e16409. <https://doi.org/10.2196/16409>
- Turing, A. M. (2012). *Computing machinery and intelligence (1950)*. *The Essential Turing: The Ideas That Gave Birth to the Computer Age*, 433-464.
- Twomey, C., O'Reilly, G., & Meyer, B. (2017). Effectiveness of an individually-tailored computerised CBT programme (Deprexis) for depression: A meta-analysis. *Psychiatry Research*, 256, 371–377. <https://doi.org/10.1016/j.psychres.2017.06.081>
- Usher, K., Durkin, J., & Bhullar, N. (2020). The COVID-19 pandemic and mental health impacts. *International Journal of Mental Health Nursing*, 29, 315–318. doi: 10.1111/inm.12726
- Vargas Gallegos, B., Gumucio Dobbs, F., & González Guevara, C. (2015). Efectividad en intervenciones terapéuticas con niños y adolescentes: Factores asociados a la Persona del Terapeuta y la Alianza Terapéutica. *Revista Chilena de Psiquiatría y Neurología de la Infancia y Adolescencia*, 24, 10-17.
- Vizard, T., Davis, J., White, E., & Beynon, B. (2020). *Coronavirus and depression in adults, Great Britain: June 2020*. Office for National Statistics.
- Wahle, F., Bollhalder, L., Kowatsch, T., & Fleisch, E. (2017). Toward the design of evidence-based mental health information systems for people with depression: A systematic literature review and meta-analysis. *Journal of Medical Internet Research*, 19(5), e191. <https://doi.org/10.2196/jmir.7381>
- Wampold, B. E., & Imel, D. E. (2021). *El gran debate de la psicoterapia: La evidencia de qué hace que la terapia funcione*. Eleftheria.
- Wang, Y., & Gorenstein, C. (2013). Psychometric properties of the Beck Depression Inventory-II: a comprehensive review. *Revista Brasileira de Psiquiatria*, 35(4), 416–431. <https://doi.org/10.1590/1516-4446-2012-1048>

- Warmerdam, L., van Straten, A., Jongasma, J., Twisk, J., & Cuijpers, P. (2010). Online cognitive behavioral therapy and problem-solving therapy for depressive symptoms: Exploring mechanisms of change. *Journal of Behavior Therapy and Experimental Psychiatry*, 41(1), 64–70. <https://doi.org/10.1016/j.jbtep.2009.10.003>.
- Watts, S., Mackenzie, A., Thomas, C., Griskaitis, A., Mewton, L., Williams, A., & Andrews, G. (2013). CBT for depression: A pilot RCT comparing mobile phone vs. computer. *BMC Psychiatry*, 13, 49. <https://doi.org/10.1186/1471-244X-13-49>. <http://www.biomedcentral.com/1471-244X/13/49>
- Weitz, E., Kleiboer, A., Van Straten, A., & Cuijpers, P. (2018). The effects of psychotherapy for depression on anxiety symptoms: A meta-analysis. *Psychological Medicine*, 48(13), 2140–2152. <https://doi.org/10.1017/S0033291717003622>.
- Wiederhold, B. K. (2020). Connecting through technology during the coronavirus disease 2019 pandemic: Avoiding “Zoom Fatigue”. *Cyberpsychology, Behavior, and Social Networking*, 23(7), 437-438. <https://doi.org/10.1089/cyber.2020.29188.bkw>

ANEXOS

Tabla A1

Variables metodológicas

Autor (año)	Metodología	Grupo control*	N**	Medida de resultados	Seguimiento	Pérdida de sujetos
<i>Berger et al. (2011)</i>	Grupo control vs 2 grupos experimentales (Deprexis)	LDE	76	BDI-II	6 meses	Sí
<i>Berger et al. (2017)</i>	Grupo control vs grupo experimental (Velibra iCBT)	TH IP	139	BDI-II	6 meses	Sí
<i>Berger et al. (2018)</i>	Grupo control vs grupo experimental (Deprexis)	TH IP	98	BDI-II GAD-7	6 meses	Sí
<i>Fitzpatrick et al. (2017)</i>	Grupo control vs grupo experimental (Woebot)	Lecturas	70	GAD-7 PHQ-9	-	Sí
<i>Fitzsimmons-Craft et al. (2021)</i>	Grupo control vs 2 grupos experimentales (Tessa – studentBodies)	LDE	700	PHQ-8 GAD-7	3 y 6 meses	Sí
<i>Gold et al. (2023)</i>	Grupo control vs 2 grupos experimentales (Amiria iCBT)	TH IP	262	BDI-II	3 meses	Sí

<i>Meyer et al. (2009)</i>	Grupo control vs grupo experimental (Deprexis)	LDE	396	BDI-II	6 meses	Sí
<i>Moritz et al. (2012)</i>	Grupo control vs grupo experimental (Deprexis)	LDE	210	BDI-II	-	Sí
<i>Oh et al. (2020)</i>	Grupo control vs grupo experimental (Mobile app-based interactive CBT)	TH Lecturas	41	HADS HADS	-	No
<i>Richards et al. (2015)</i>	Grupo control vs grupo experimental (Computerized cognitive-behavior therapy (iCBT))	LDE	188	BDI-II GAD-7	3 y 6 meses	Sí
<i>Sadeh et al. (2023)</i>	Grupo control vs grupo experimental (Eleos health platform)	TH IP	47	PHQ-9 GAD-7	-	Sí
<i>Sandoval et al. (2017)</i>	Grupo control vs grupo experimental (imbPST)	NT	45	BDI-II	2 meses	No
<i>Schroder et al. (2014)</i>	Grupo control vs grupo experimental (Deprexis)	LDE	78	BDI-II	6 meses	Sí
<i>Zwerenz et al. (2017)</i>	Grupo control vs grupo experimental (Deprexis)	TH IP	229	BDI-II PHQ-9 GAD-7	-	Sí

Nota: todos los estudios seleccionados son ensayos clínicos aleatorizados (ECA)

*= LDE = Lista de espera; TH = Tratamiento habitual; NT = No tratamiento; IP = Intervención personal (cara a cara)

** = Número total de la muestra

Tabla A2

Variables de la muestra

ESTUDIO	MUESTRA	SEXO		EDAD (MEDIA ± DS)*	MUESTRA**	PAÍS	RECLUTAMIENTO
		M	F				
<i>Berger et al. (2011)</i>	P. clínica Diagnóstico dep	30.3%	69.7%	38.8 ± 14.0	T: 76 I: 50 / C: 26	Alemania y Suiza	Internet
<i>Berger et al. (2017)</i>	P. clínica Tr. de ansiedad	29.5%	70.5%	41,95 ± 12,2	T: 139 I: 70 / C: 69	Alemania, Suiza y Austria	Internet Cara a cara
<i>Berger et al. (2018)</i>	P. clínica Tr. afectivo	33.6%	66.4%	43,05 ± 12,1	T: 98 I: 51 / C: 47	Alemania	Cara a cara – pacientes ambulatorios
<i>Fitzpatrick et al. (2017)</i>	P. no clínica S dep y s ans	19%	81%	22,20 ± 2,31	T: 70 I: 34 / C: 36	Estados Unidos	Internet
<i>Fitzsimmons- Craft et al. (2021)</i>	P. clínica Tr. alimentario	-	100%	21,08 ± 3,09	T: 700 I: 352 / C: 348	Estados Unidos	Internet
<i>Gold et al. (2023)</i>	P. clínica Esclerosis múltiple	73.7%	76.3%	46,96 ± 11,7	T: 262 I: 174 / C: 88	Alemania y Estados Unidos	Cara a cara – pacientes ambulatorios

<i>Meyer et al. (2009)</i>	P. no clínica S dep	24%	76%	34,76 ± 11,60	T: 396 I: 320 / C: 76	Alemania	Internet
<i>Moritz et al. (2012)</i>	P. clínica S dep	21.5%	78.5%	38,56 ± 13,29	T: 210 I: 105 / C: 105	Alemania	Internet
<i>Oh et al. (2020)</i>	P. clínica Tr. de pánico	48.8%	51.2%	41,05 ± 11,55	T: 41 I: 21 / C: 20	Corea del sur	Cara a cara – pacientes ambulatorios
<i>Richards et al. (2015)</i>	P. no clínica S dep	27%	73%	39,86 ± 10,94	T: 188 I: 96 / C: 92	Irlanda	Cara a cara
<i>Sadeh et al. (2023)</i>	P. clínica S dep y s ans	27.7%	72.3%	30,59 ± 10,76	T: 47 I: 23 / C: 24	Estados Unidos	Internet
<i>Sandoval et al. (2017)</i>	P. clínica Tr. depresivo mayor o distimia	37.8%	62.2%	28,62 ± 9,14	T: 45 I: 25 / C: 20	Estados Unidos	Internet
<i>Schroder et al. (2014)</i>	P. clínica Epilepsia y s dep	24.4%	75.6%	37,53 ± 10,92	T: 78 I: 38 / C: 40	Alemania	Email y online fórums
<i>Zwerenz et al. (2017)</i>	P. clínica S dep	39.3%	60.7%	47.98 ± 9.79	T: 229 I: 115 / C: 114	Alemania	Internet

P. = población; Dep = Depresiva ; Ans = Ansiosa ; S = sintomatología; Tr. = trastorno

* = Edad media \pm desviaciones típicas

** = Tamaño de la muestra; T = Total; I = grupo intervención; C = grupo control