



UNIVERSIDAD PERUANA
CAYETANO HEREDIA

Facultad de
MEDICINA

CONCORDANCIA DIAGNÓSTICA ENTRE 4 DIFERENTES FORMAS DE
DEFINIR DEPRESIÓN (PHQ-2, PHQ -9, DSM-5 Y CIE-10) EN UNA
POBLACIÓN ADULTA BASADA EN LA ENDES 2019 -PERÚ

DIAGNOSTIC CONCORDANCE BETWEEN 4 DIFFERENT WAYS OF
DEFINING DEPRESSION (PHQ-2, PHQ-9, DSM-5 AND ICD-10) IN AN
ADULT POPULATION BASED ON ENDES 2019 -PERU

TRABAJO DE INVESTIGACIÓN PARA OPTAR POR EL TÍTULO
PROFESIONAL DE MÉDICO CIRUJANO

ALUMNOS:

PAVEL MORI VILLALVA

FABIOLA STEPHANY MURICHES FÉLIX

ASESOR:

JOHANN MARTÍN VEGA DIENSTMAIER

LIMA - PERÚ

2022

JURADO

Presidente: Dra. Silvana Virginia Sarabia Arce

Vocal: Dra. Paola Albinagorta Ariza

Secretario: Dr. Edward Antonio Lozano Vargas

Fecha de sustentación: 21 de mayo del 2022

Calificación: Aprobado

ASESOR DEL TRABAJO DE INVESTIGACIÓN

ASESOR

Dr. Johann Martín Vega Dienstmaier

Departamento de Clínicas Médicas

ORCID: 0000-0002-5686-4014

DEDICATORIA

A Dios por habernos guiado a lo largo de nuestras vidas y a nuestras familias, en especial a nuestros padres que fueron nuestro soporte y fuente de inspiración.

AGRADECIMIENTOS

Al Dr. Johann Vega Dienstmaier por su apoyo y motivación constante.

FUENTES DE FINANCIAMIENTO

Autofinanciado

DECLARACIÓN DE CONFLICTO DE INTERÉS

Los autores declaran no tener conflictos de interés

TABLA DE CONTENIDOS

	Pág.
I. INTRODUCCIÓN	1
II. OBJETIVOS	7
III. MATERIALES Y MÉTODOS	8
IV. RESULTADOS	11
V. DISCUSIÓN	14
VI. CONCLUSIONES	18
VII. REFERENCIAS BIBLIOGRÁFICAS	19
VIII. TABLAS	25

RESUMEN

Antecedentes: Existen muchos métodos para detectar y definir depresión, que idealmente deberían dar los mismos resultados. **Objetivos:** Determinar la correspondencia entre 4 formas de definir depresión a partir de la información que proporciona el PHQ-9: utilizando los puntos de corte del PHQ-9 y del PHQ-2 y los criterios del DSM-5 y de la CIE-10. **Materiales y métodos:** Se realizó un análisis secundario de la base de datos de la ENDES 2019 que utilizó una muestra probabilística de 31726 adultos y adultos mayores. Se evaluó la concordancia (índice kappa) entre las diferentes formas de definir depresión y se determinó su prevalencia de acuerdo a cada una de estas formas. **Resultados:** Los coeficientes kappa entre las formas de definir depresión variaron desde 0,24 (entre el PHQ-2 y el DSM-5) hasta 0,84 (entre el DSM-5 y la CIE 10). Las prevalencias de depresión fueron 10,6%, 6,8%, 2,6%, 1,6% y 2,16% de acuerdo a las definiciones correspondientes a PHQ-2, PHQ-9 ≥ 10 , PHQ-9 ≥ 15 , DSM-5 y CIE-10, respectivamente. **Conclusiones:** La prevalencia de depresión varía según el método que se emplee para definirla. Para el PHQ-9, el punto de corte de 15 concuerda más que el de 10 con los criterios de depresión del DSM-5 y la CIE-10. Al utilizar el DSM-5 se observaron menores prevalencias de depresión que al emplear la CIE-10.

Palabras claves: Depresión, PHQ-9, PHQ-2, DSM-5, CIE-10 (Clasificación Internacional de Enfermedades)

ABSTRACT

Background: There are many methods to detect and define depression, which ideally should give the same results. **Objectives:** To determine the correspondence between 4 ways of defining depression based on the information provided by the PHQ-9: using the PHQ-9 and PHQ-2 cut-off points and the DSM-5 and ICD-10 criteria. **Materials and methods:** A secondary analysis of the ENDES 2019 database was carried out using a probabilistic sample of 31,726 adults and older adults. The concordance (kappa index) between the different ways of defining depression was evaluated and its prevalence was determined according to each of these ways. **Results:** The kappa coefficients between the ways of defining depression ranged from 0.24 (between PHQ-2 and DSM-5) to 0.84 (between DSM-5 and ICD-10). The prevalence showed the following variation 10.6%, 6.8%, 2.6%, 1.6% and 2.16% corresponding to the PHQ-2, PHQ-9 >10, PHQ-9 >15 scales, DSM-5 and ICD-10, respectively. **Conclusions:** The prevalence of depression varies according to the method used to define it. For the PHQ-9, the cut-off point of 15 agrees more with the DSM-5 and ICD-10 criteria of depression than that of 10. When using the DSM-5, lower prevalence of depression were observed than when using the ICD-10.

Keywords: Depression, PHQ-9, PHQ-2, DSM-5, ICD-10 (International Classification of Diseases)

I. INTRODUCCIÓN

Los trastornos mentales más frecuentes en la población peruana son la ansiedad y la depresión (1,2,3) y tienen un gran impacto no solo en la persona que los padece sino también en sus familiares y la población en general. La Organización Mundial de la Salud (OMS) considera la depresión como la primera causa de discapacidad a nivel mundial, afectando a más de 280 millones de personas (3,8% de la población) (4,5). La prevalencia de vida en el trastorno depresivo mayor a nivel mundial varía según el grupo etario y alcanza su pico máximo en la edad más adulta (niveles superiores de 7,5% en mujeres con un rango de edad entre los 55 y 74 años, y por encima de 5,5% en hombres). La depresión también ocurre en niños y adolescentes menores de 15 años, pero con tasas de prevalencia por debajo de los grupos de mayor edad (6).

Alrededor del mundo, la prevalencia anual reportada del trastorno depresivo mayor de cada país varía desde un 3% reportado en Japón hasta un 16,9% reportado en Estados Unidos de América; para los demás países en general la prevalencia fluctúa en un rango de 8-12% (7). En el Perú, el Instituto Nacional de Salud Mental "Honorio Delgado - Hideyo Noguchi" realizó un estudio epidemiológico en el año 2012, el cual evidencia que la prevalencia de vida de depresión en Lima Metropolitana y Callao es 17,3%, siendo mayor en las mujeres (21,6%) que en los varones (12,6%) (8).

La presencia de los síntomas caracterizados fundamentalmente por humor depresivo o tristeza, pérdida del interés o del disfrute (anhedonia) y cansancio o

fatiga (9) genera una discapacidad directa en los individuos, afectando su calidad de vida, además hay una importante estigmatización ejercida por la sociedad que conlleva a la discriminación (10).

Al ser el trastorno depresivo mayor una enfermedad que genera discapacidad en ambos sexos y diversos grupos etarios, es necesario que el tamizaje se pueda realizar en diferentes grupos poblacionales y de esta manera detectar los casos que aún no hayan sido diagnosticados y así poder brindarles tratamiento. Por esta razón, las pruebas de detección temprana o screening para el trastorno depresivo mayor no solo tienen que ser lo suficientemente fiables para ser utilizadas en los centros de atención primaria sino también de fácil acceso y aplicación. Posteriormente se debe confirmar los casos probables mediante las entrevistas clínicas, brindando el correcto diagnóstico y tratamiento al paciente (11).

Una de las escalas de tamizaje que podría usarse en centros de atención primaria para identificar al trastorno depresivo mayor es el Patient Health Questionnaire (PHQ-9) que es parte de la PRIME-MD (Primary Care Evaluation of Mental Disorders) que originalmente se creó para identificar 5 principales desórdenes mentales: depresión, ansiedad, abuso de alcohol, desórdenes somatomorfos y desórdenes alimenticios. Diversos estudios han evaluado el PHQ-9 como prueba de detección, siendo el valor de 10 el punto de corte más recomendado, proporcionando la mejor combinación de sensibilidad (0,88) y especificidad (0,88) (12).

Gilbody et al encontraron una sensibilidad de 0,80 y una especificidad 0,92 respecto al *gold standard* DSM-IV, al evaluar el método por algoritmo diagnóstico, el cual requiere 5 o más de los 9 criterios que definen depresión siempre que, uno de los síntomas sea estado de ánimo deprimido o anhedonia y que se presente “más de la mitad de los días” en las últimas dos semanas; por el contrario, cuando se usa el punto de corte mayor o igual a 10 se encuentra una sensibilidad y especificidad ambas de 0,88 (12,13). Se realizó una revisión sistemática de 17 estudios, en la cual se evaluó diversas herramientas para la identificación del trastorno depresivo mayor, concluyendo que el PHQ-9 también es una medida confiable y válida para evaluar la gravedad de la depresión, lo cual hace que sea una buena herramienta clínica y de investigación (12). El estudio anteriormente mencionado junto con Manea et al, ambos meta-análisis, encontraron buenas propiedades psicométricas del PHQ-9 para la detección del trastorno depresivo mayor (14,15).

Planteamiento del problema:

¿Cuál es la correspondencia entre 4 formas de definir depresión a partir del PHQ-9: utilizando los puntos de corte del PHQ-9 y del PHQ-2; y los criterios del DSM-5 y de la CIE-10?

Justificación:

Según la opinión de expertos peruanos se concluyó que, el PHQ-9 podría ser una gran herramienta válida para el tamizaje del trastorno depresivo mayor y puede ser aplicado en el Perú (16). Así, este trabajo podría servir como base para futuras investigaciones que necesiten utilizar el PHQ-9, con lo que esperamos brindar

información valiosa sobre cómo puede variar la detección de la depresión y la estimación de su prevalencia utilizando las distintas formas o definiciones para el diagnóstico de este trastorno en la población peruana.

Marco teórico:

El PHQ-9 evalúa la presencia de los nueve síntomas especificados en la 5ª edición del Manual Diagnóstico y Estadístico de los Trastornos Mentales (DSM-5), durante al menos un periodo de 2 semanas, los cuales son: estado de ánimo deprimido, anhedonia, problemas de sueño, cansancio o fatiga, cambios en el apetito o el peso, sentimientos de culpa o inutilidad, dificultad para concentrarse, lentitud o inquietud, e ideación suicida. Dependiendo de la frecuencia con respecto a los días ("nunca", "algunos días", "más de la mitad de los días", y "casi todos los días") se puede obtener entre 0 a 3 puntos por cada ítem (12).

Con respecto a los criterios del DSM-5 para el diagnóstico del trastorno depresivo mayor se define como la presencia de cinco (o más) de los siguientes 9 síntomas durante un período no menor de 2 semanas durante la mayor parte del tiempo: estado de ánimo depresivo, disminución del interés o de la capacidad para el placer en todas o casi todas las actividades, pérdida o aumento importante de peso, pérdida o aumento del apetito, insomnio o hipersomnia, agitación o enlentecimiento en la psicomotricidad, fatiga o pérdida de energía, sentimientos de inutilidad o de culpa de forma excesiva, disminución de la capacidad para concentrarse o pensar, pensamientos frecuentes de muerte o ideación suicida; es importante tener en cuenta

que para establecer el diagnóstico uno de los síntomas necesariamente debe ser estado de ánimo depresivo o la anhedonia (17).

La Clasificación Internacional de Enfermedades (CIE-10), para el diagnóstico definitivo del episodio depresivo leve, requiere 2 de estos 3 síntomas: ánimo depresivo, pérdida de interés y de la capacidad de disfrutar (anhedonia), y aumento de la fatigabilidad, además de al menos uno de los siguientes síntomas: disminución de la atención y concentración, pérdida de la confianza en sí mismo y sentirse inferior, ideas de culpa, pensamientos y actos suicidas o de autoagresiones, agitación, trastornos del sueño y pérdida del apetito (18).

El PHQ-2 consta de las primeras 2 preguntas del PHQ-9, que se refieren a estado de ánimo deprimido y anhedonia evaluados en un periodo de 2 semanas. Dependiendo de la frecuencia con respecto a los días ("nunca", "algunos días", "más de la mitad de los días" y "casi todos los días") se puede obtener entre 0 a 3 puntos por cada ítem. Las puntuaciones oscilan entre 0 y 6 y el punto de corte recomendado es de 3 o más. Las acciones recomendadas para las personas que puntúan 3 o más son una de las siguientes: administrar el PHQ-9 completo y realizar una entrevista clínica para evaluar el trastorno depresivo mayor (19,20).

En un metanálisis publicado en el 2020 se encontró que para el PHQ-2 con un punto de corte > 3 , la especificidad y la sensibilidad para la detección de Trastorno depresivo mayor fueron 0,85 y 0,72 respectivamente. Sin embargo, se evidenció que al combinar el PHQ-2 con el PHQ-9 (con punto de corte ≥ 10) se obtuvo una

sensibilidad semejante pero una especificidad superior comparada con el PHQ-9 únicamente (21).

El PHQ-9 es un instrumento que evalúa los síntomas depresivos en base a los cuales se puede hacer el diagnóstico de depresión siguiendo los criterios del DSM-5, asimismo, también recoge la mayor parte de la sintomatología necesaria para hacer el diagnóstico de depresión en base al CIE-10. Por otro lado, el PHQ-2 es un cribado inicial que se basa en 2 preguntas: “¿Ha tenido falta de interés o poca satisfacción en hacer cosas?” y “¿se ha sentido decaído, deprimido o desesperanzado?”, las cuales forman parte del cuestionario PHQ-9 (19,20).

El PHQ-9 y el PHQ-2 nos permiten definir casos probables de depresión utilizando los puntos de corte. Por lo expuesto, se puede definir depresión de 4 formas: Cuando el PHQ-9 alcanza o supera un puntaje (punto de corte mayor o igual a 10 o 15), en base a un punto de corte del PHQ-2 (punto de corte mayor o igual a 3), en base a los criterios diagnósticos de depresión mayor del DSM-5 y en base a los criterios diagnóstico de la CIE-10 (19,20).

II. OBJETIVOS

General:

Determinar la correspondencia entre 4 formas de definir depresión a partir de la información que proporciona el PHQ-9: utilizando los puntos de corte del PHQ-9 y del PHQ-2; y los criterios del DSM-5 y de la CIE-10.

Específicos:

- Precisar el punto de corte del PHQ-9 (≥ 10 o ≥ 15) que más se asocia al diagnóstico de depresión mayor basado en los criterios del DSM-5 y CIE-10.
- Estimar la prevalencia de depresión según criterios del PHQ-2, PHQ-9, DSM-5 y CIE-10 según edad (< 65 o ≥ 65) y sexo (femenino o masculino) en nuestra población.
- Determinar el grado de concordancia entre el diagnóstico de depresión basado en los puntos de corte del PHQ-9, PHQ-2 y los criterios de la CIE-10 y el DSM-5, en general, por edad (< 65 o ≥ 65) y sexo (femenino o masculino) en nuestra población.

III. MATERIALES Y MÉTODOS

Diseño del estudio:

Se realizó un análisis secundario de la base de datos de la ENDES 2019, la cual proviene de un estudio observacional de tipo transversal.

Población y muestra:

El muestreo realizado por el INEI fue bietápico, probabilístico, estratificado, de tipo auto ponderado y diseñado para brindar estimaciones representativas de la población. La muestra de estudio consistió en 31726 adultos de ambos sexos de 18 años a más entrevistados el año 2019, residentes de las zonas rurales o urbanas del Perú.

Criterios de elegibilidad:

Criterios de inclusión:

- Participantes que tengan de 18 años a más al momento de realizar la encuesta ENDES 2019.
- Participantes que hayan contestado la pregunta 700 del Cuestionario de Salud de la ENDES que contiene las nueve sub-preguntas (A - I) del Instrumento PHQ-9 a partir de las cuales se construyen las variables relativas a depresión.

Criterios de exclusión:

- Participantes que no cumplan los criterios de selección.

Variables:

La definición operacional de cada variable se encuentra detallada en la tabla 1.

Procedimientos y técnicas:

La base de datos de La Encuesta Demográfica y de Salud Familiar (ENDES) del año 2019 incluye una sección sobre salud mental, y dentro de ésta, el Cuestionario PHQ-9.

Las variables utilizadas en el estudio corresponden a la pregunta 700 del Cuestionario de Salud de la ENDES. Dicha pregunta contiene nueve sub - preguntas del instrumento PHQ 9 (700 -A a 700 -I) las cuales son analizadas según la escala que va de 0 a 3 puntos, con una puntuación total de 0 a 27 puntos. Las variables correspondientes a las diferentes formas de definir depresión se construyeron de la manera especificada en la Tabla 1.

Para seleccionar el punto de corte alternativo se eligió aquél que según diversas fuentes bibliográficas mostró mejores propiedades psicométricas. En el estudio se probaron dos puntos de corte ($\text{PHQ-9} \geq 10$ y $\text{PHQ-9} \geq 15$) utilizando como referencia tanto el DSM-5 como el CIE-10.

Consideraciones éticas:

El proyecto de investigación fue aprobado por el Comité Institucional de Ética en Investigación de la Universidad Peruana Cayetano Heredia. El presente trabajo es un estudio de fuente secundaria: la base de datos de la ENDES, la cual es de uso libre y se encuentra disponible en la página web del Instituto Nacional de

Estadística e Informática (INEI) del Perú; y carece de la información necesaria para identificar a los participantes.

Análisis de datos:

Se calculó el índice Kappa entre cada una de las siguientes formas de definir depresión: PHQ-9 \geq 10, PHQ-9 \geq 15, PHQ-2 \geq 3, criterios del DSM-5 y criterios de la CIE-10. Para el cálculo de la prevalencia se utilizaron los pesos estadísticos que se encuentran incluidos en la base de datos de la ENDES 2019. Los análisis se realizaron para la muestra total y por grupos según sexo (hombres y mujeres) y edad (<65 y \geq 65 años). Se utilizó el programa estadístico STATA 17.

Para el grado de concordancia o acuerdo entre las diferentes formas de diagnóstico, se usó el coeficiente kappa, el cual se interpreta de la siguiente manera: un kappa < 0 significa que el acuerdo es inexistente; entre 0,01 y 0,2, es leve o insignificante; entre 0,21 y 0,4, es aceptable; entre 0,41 y 0,60, es moderado; entre 0,61 y 0,8, es sustancial o considerable; y entre 0,81 y 1,00, es casi perfecto (22).

IV. RESULTADOS

Descripción de la muestra

La muestra estuvo conformada por 31726 individuos, 57,71% de mujeres y 42,29% de varones, lo cual corresponde a 51,86% (IC-95%: 50,94%-52,78%) de mujeres y 48,14% (IC-95%: 47,22%-49,06%) de varones en la población.

La edad promedio en la muestra fue 41,33 años (DE=16,66) en la muestra, correspondiendo a 42,95 años (IC-95%: 42,62 a 43,29) en la población. La muestra estaba compuesta por 11,81% de personas de 65 años o más: 13,23% en la población (IC-95%: 12,6% - 13,9%).

Concordancia entre las diferentes formas de definir depresión

En la tabla 2 se muestra el grado de concordancia entre los métodos utilizados para definir depresión (PHQ-2, PHQ-9 ≥ 10 , PHQ-9 ≥ 15 , DSM-5 y CIE-10) para el total de la muestra y para los subgrupos de acuerdo a sexo y edad. Los resultados muestran concordancias variables, pero todas significativas (valor $p < 0,0001$).

Entre las definiciones de depresión, se obtuvo una fuerza de concordancia interpretada como aceptable para la relación PHQ-2 y PHQ-9 ≥ 15 , PHQ-2 y DSM-5, PHQ-2 y CIE 10, y PHQ-9 y DSM-5. Asimismo, se observó que la relación entre el PHQ-9 ≥ 10 y PHQ-9 ≥ 15 alcanza un nivel moderado según el índice de concordancia, así como la relación entre el PHQ-9 ≥ 10 y la CIE-10.

Con respecto, al coeficiente Kappa entre los instrumentos PHQ-2 y PHQ-9 ≥ 10 , PHQ-9 ≥ 15 y CIE 10, y PHQ-9 ≥ 15 y DSM-5 es considerado como considerable o sustancial. Se obtiene así en todas las relaciones antes mencionadas un menor grado de concordancia comparada con la valoración casi perfecta observada entre los criterios del DSM-5 y la CIE 10.

En el sexo femenino, el grado de concordancia más alto fue el de la relación entre el DSM-5 y la CIE 10; y el menor, con una concordancia aceptable, fue entre el PHQ-2 y el DSM-5. Por otro lado, en lo que respecta al sexo masculino, los valores de Kappa para las mismas relaciones entre DSM-5 y CIE10, al igual que entre PHQ-2 y DSM-5, mostraron similar grado de coincidencia. La fuerza de concordancia entre las definiciones de depresión fue más alta en la población mayor o igual a los 65 años de edad.

Los resultados para ambos sistemas de clasificación (DSM-5 y CIE-10) muestran que el mejor punto de corte para el PHQ-9 es ≥ 15 con una prevalencia estimada para el diagnóstico de depresión de 2,6% (IC95% 2,34-2,87) el cual, se observa en la tabla 3.

Prevalencias de depresión según las diferentes formas de definir su presencia

En la tabla 3 se evidencia que la prevalencia obtenida con los diversos métodos es 10,6%, 6,8%, 2,6%, 1,6% y 2,16% correspondiente al PHQ-2, PHQ-9 ≥ 10 , PHQ-9 ≥ 15 , DSM-5 y CIE-10, respectivamente.

Dentro de la población de estudio se encontró que la mayor prevalencia encontrada en el sexo femenino fue al emplear el PHQ-2 (13,7%), además se observó que al utilizar tanto las escalas como los instrumentos de referencia los valores de prevalencia en el sexo femenino son mayores que los obtenidos para el sexo masculino. Al evaluar la variable edad se halló que la población con edad ≥ 65 años, alcanzó valores de prevalencia superiores a los encontrados en < 65 años.

V. DISCUSIÓN

En este estudio se evaluó la concordancia entre diferentes métodos para definir depresión y su prevalencia en base a ellos en nuestra población; tanto para los participantes en general como para los grupos basados en sexo y edad.

La prevalencia ponderada de depresión, para cada una de sus definiciones, es significativamente mayor en el sexo femenino y en ≥ 65 años. Además se encontró que la prevalencia en la población general al usar los diferentes métodos diagnósticos presenta una gran variación, siendo las escalas (PHQ-2 y PHQ-9) las que clasifican a más individuos como deprimidos en comparación con los sistemas diagnósticos (DSM-5 y CIE-10). Hay evidencia en la literatura de que diferentes formas de establecer el diagnóstico de depresión no son equivalentes, por ejemplo, Levis B et al, evaluaron 3 diferentes entrevistas diagnósticas [Structured Clinical Interview for DSM (SCID), Mini International Neuropsychiatric Interview (MINI) and Composite International Diagnostic Interview (CIDI)], y concluyeron que al ser comparadas no clasifican de manera similar la depresión mayor (23).

Con respecto a la prevalencia ponderada según el instrumento PHQ-9, al utilizar el punto de corte 15 se obtuvieron resultados más cercanos a los hallados con los sistemas de clasificación de referencia (CIE 10 y DSM-5). El punto de corte 10 generó una sobrestimación de la prevalencia de depresión respecto a la obtenida con el punto de corte 15. Lo cual es reforzado por el estudio realizado por Villareal-Zegarra en el 2020, en el cual los resultados obtenidos en el periodo de tiempo 2014-2018 son muy similares a los obtenidos en nuestro estudio (2,3% - 3%) (24).

Es ampliamente conocido que el punto de corte más utilizado para el instrumento PHQ-9 es un puntaje de 10 esto se ve representado en cinco estudios llevados a cabo en Honduras, Italia, Alemania, Arabia Saudita y Reino Unido en los cuales se encontró que la versión del PHQ-9 con un punto de corte de 10, obtuvo una adecuada especificidad y sensibilidad en relación al DSM-IV. (24,25,26,27,28) Los estudios de investigación sobre el uso del PHQ-9 en Perú son pocos, no obstante el estudio realizado por Villarreal - Zegarra en el 2020, sugiere que el puntaje del punto de corte mayor igual a 15 para el tamizaje de depresión puede tener propiedades diagnósticas aceptables, ya que a partir de este valor se clasifica la depresión como moderada-severa según el algoritmo diagnóstico del PHQ-9, con valores de sensibilidad de 68% y una especificidad de 95% comparado con el DSM-5, la cual fue el patrón de referencia en dicho estudio (12,24).

Adicionalmente, en otros estudios el punto de corte mayor o igual a 15 es de suma importancia ya que define si un episodio depresivo mayor requiere o no tratamiento (24). Estos hallazgos apoyan los resultados encontrados en el presente estudio, en el que se sugiere como punto de corte óptimo al puntaje de 15 sobre la puntuación de 10. Además, en Perú se encontró valores similares de prevalencia al comparar el obtenido con el PHQ-9 (2,6%) con el World Mental Health en el que se encontró un valor de 2,7% para el Trastorno de depresión mayor en el 2005 (30).

En nuestro estudio se encontró un grado de acuerdo sustancial entre el PHQ-2 y PHQ-9 ≥ 10 , PHQ-9 ≥ 15 y DSM-5, y PHQ-9 ≥ 15 y CIE-10. El mayor kappa encontrado fue entre el DSM-5 y la CIE-10 que corresponde a un acuerdo “casi perfecto”.

Aunque la concordancia entre la CIE 10 y el DSM-5 es alta, se observa que el umbral diagnóstico de la CIE es más bajo que el del DSM-5, lo cual se refleja en menores prevalencias al usarse este último. Según los resultados obtenidos se propone que existe una discordancia en la prevalencia entre los dos sistemas diagnósticos, ya que para el CIE 10 es indispensable tener al menos 4 síntomas a diferencia de 5 síntomas depresivos en el caso del DSM-5, esto podría generar en la práctica clínica una sobreestimación al usar los criterios de la CIE-10 sobre los del DSM-5. Saito M et al, concluye que esta diferencia se debe a que la CIE-10 es más sensible al diagnosticar los episodios leves de depresión a comparación del DSM-IV que presentó una mayor sensibilidad para los episodios moderados a graves (31).

La fuerza de concordancia al utilizar las escalas PHQ-2, $\text{PHQ-9} \geq 15$ y $\text{PHQ-9} \geq 10$ tiende a ser mayor en el grupo mayor igual de 65 años. Matias A et al, encontró que en pacientes con una edad mayor o igual de 60 años, el coeficiente de confiabilidad Kappa fue moderado con un nivel de significancia de $p < 0,001$ al comparar el PHQ-9 y la Escala de Depresión Geriátrica de 15 ítems, lo cual es coherente con nuestro estudio (32).

Se presentaron limitaciones con respecto a algunas variables que forman parte de los criterios diagnósticos del CIE-10 que no fueron consideradas en el PHQ-9, específicamente, tales como disminución marcada de la libido y perspectiva sombría del futuro, otra limitación fue que la detección de depresión no fue respaldada con una evaluación psiquiátrica, por ende, solo nos basamos en los síntomas depresivos recogidos por los ítems del PHQ-9 para determinar los casos.

También es importante incluir las limitaciones propias de utilizar un análisis secundario como la ENDES (Encuesta Demográfica y de Salud Familiar) una de ellas es que al no ser realizada específicamente para la detección de depresión incluye otros cuestionarios lo cual podría influir de forma negativa produciendo fatiga por la extensión de la entrevista, generando así una mala disposición para responder las preguntas adecuadamente pese a que se cuente con un personal capacitado para la entrevista directa y recojo de la información (33). Debido a que no se realizó una entrevista por un psiquiatra o algún profesional de la salud no se pudo excluir que la causa de depresión se deba a un trastorno bipolar, trastorno esquizoafectivo, esquizofrenia, trastorno esquizofreniforme, trastorno delirante u otro trastorno especificado o psicótico, consumo de sustancias u otra afección médica.

VI. CONCLUSIONES

Los hallazgos del presente estudio muestran que la prevalencia de depresión varía ampliamente según el método que se emplee para definirla, desde 1,6% de acuerdo al DSM-5 hasta 10,6% según el PHQ-2. Asimismo, la definición de depresión basada en el PHQ-9 que concuerda más con las definiciones según la CIE-10 y el DSM es la que usa el punto de corte de ≥ 15 en la población de estudio. Aunque la concordancia entre la CIE-10 y DSM-5 es alta, se observa que el umbral diagnóstico de la CIE-10 es más bajo que el del DSM-5, lo cual se refleja en menores prevalencias al usarse este último.

VII. REFERENCIAS BIBLIOGRÁFICAS

1. Instituto Especializado de Salud Mental “Honorio Delgado – Hideyo Noguchi”. Estudio epidemiológico metropolitano en salud mental 2002. Informe general. Lima: IESM HD-HN; 2002.
2. Instituto Especializado de Salud Mental «Honorio Delgado – Hideyo Noguchi». Estudio Epidemiológico de Salud Mental en Huaraz 2003. Informe General. Lima, 2004.
3. Instituto Especializado de Salud Mental. Estudio Epidemiológico en Salud Mental en la Selva Peruana 2004. Informe General. Anales de Salud Mental Vol. XXI. Año 2005, Número 1 y 2. Lima.
4. World Health Organization. Mental health: Depression. Fact Sheets, Detail, Depression revised September 2021 [página de internet]. Fecha de acceso: Febrero 2022. Disponible en: <https://www.who.int/es/news-room/fact-sheets/detail/depression>
5. Schulz PE, Arora G. Depression. CONTINUUM: Lifelong Learning in Neurology. 2015;21(3):756-771.
6. Organización Panamericana de Salud (OPS). Depresión y otros trastornos mentales comunes. Estimaciones Sanitarias Mundiales. Washington, D.C.: Organización Panamericana de Salud, 2017 [Consultado el 5 febrero 2021]. Disponible en:

<https://iris.paho.org/bitstream/handle/10665.2/34006/PAHONMH17005-spa.pdf?sequence=1&isAllowed=y>

7. Kessler R, Berglund P, Demler O, Jin R, Koretz D, Merikangas K et al. The Epidemiology of Major Depressive Disorder. *JAMA*. 2003;289(23):3095.
8. Instituto Nacional de Salud Mental. Estudio Epidemiológico de Salud Mental en Lima Metropolitana y Callao Replicación 2012. Informe General. *Anales de Salud Mental Vol. XXIX*. Año 2013, Suplemento 1. Lima
9. MINSA. Guía de Práctica Clínica en Trastornos Mentales y del Comportamiento debido al consumo de sustancias psicotropas. Lima: Dirección General de Salud de las Personas. Dirección de Salud Mental; 2006.
10. Puentes-Rosas, Esteban, Medina-Mora, María Elena, Lozano, Rafael, Belló, Mariana, Prevalencia y diagnóstico de depresión en población adulta en México. *Salud Pública de México [Internet]*. 2005;47(1):S4-S11. Recuperado de: <https://www.redalyc.org/articulo.oa?id=10609302>
11. Malpass A, Dowrick C, Gilbody S, Robinson J, Wiles N, Duffy L et al. Usefulness of PHQ-9 in primary care to determine meaningful symptoms of low mood: a qualitative study. *British Journal of General Practice*. 2016;66(643):e78-e84.
12. Kroenke K, Spitzer R, Williams J. The PHQ-9. *Journal of General Internal Medicine*. 2001;16(9):606-613.

13. Gilbody S, Richards D, Brealey S, Hewitt C. Screening for Depression in Medical Settings with the Patient Health Questionnaire (PHQ): A Diagnostic Meta-Analysis. *Journal of General Internal Medicine*. 2007;22(11):1596-1602.
14. Williams J, Pignone M, Ramirez G, Perez Stellato C. Identifying depression in primary care: a literature synthesis of case-finding instruments. *General Hospital Psychiatry*. 2002;24(4):225-237.
15. Manea L, Gilbody S, McMillan D. Optimal cut-off score for diagnosing depression with the Patient Health Questionnaire (PHQ-9): a meta-analysis. *Canadian Medical Association Journal*. 2011;184(3): E191-E196.
16. Calderón M, Gálvez-Buccollini J, Cueva G, Ordoñez C, Bromley C, Fiestas F. Validación de la versión peruana del PHQ-9 para el diagnóstico de depresión. *Revista Peruana de Medicina Experimental y Salud Pública*. 2012;29(4):578-579.
17. Asociación Americana de Psiquiatría. Manual de diagnóstico y estadístico de los trastornos mentales(DSM-5). 5ª ed. Arlington, VA, Asociación Americana de Psiquiatría;2014.
18. World Health Organization. (2000) . Guía de bolsillo de la clasificación CIE-10: clasificación de los trastornos mentales y del comportamiento. Editorial Médica Panamericana.
19. Kroenke K, Spitzer R, Williams J. The Patient Health Questionnaire-2. *Medical Care*. 2003;41(11):1284-1292.

20. Kroenke K, Spitzer R, Williams J, Löwe B. The Patient Health Questionnaire Somatic, Anxiety, and Depressive Symptom Scales: a systematic review. *General Hospital Psychiatry*. 2010;32(4):345-359.
21. Levis B, Sun Y, He C, Wu Y, Krishnan A, Bhandari P et al. Accuracy of the PHQ-2 Alone and in Combination With the PHQ-9 for Screening to Detect Major Depression. *JAMA*. 2020;323(22):2290.
22. Abaira V. El índice kappa. *SEMERGEN - Medicina de Familia*. 2001;27(5):247-249.
23. Levis B, McMillan D, Sun Y, He C, Rice D, Krishnan A et al. Comparison of major depression diagnostic classification probability using the SCID, CIDI, and MINI diagnostic interviews among women in pregnancy or postpartum: An individual participant data meta-analysis. *International Journal of Methods in Psychiatric Research*. 2019;28(4):1-11.
24. Villarreal-Zegarra D, Cabrera-Alva M, Carrillo-Larco R, Bernabe-Ortiz A. Trends in the prevalence and treatment of depressive symptoms in Peru: a population-based study. *BMJ Open*. 2020;10(7):e036777.
25. Wulsin L, Somoza E, Heck J. The Feasibility of Using the Spanish PHQ-9 to Screen for Depression in Primary Care in Honduras. *The Primary Care Companion to The Journal of Clinical Psychiatry*. 2002;04(05):191-195.
26. Picardi A, Adler D, Abeni D, Chang H, Pasquini P, Rogers W et al. Screening for Depressive Disorders in Patients with Skin Diseases: A

- Comparison of Three Screeners. *Acta Dermato-Venereologica*. 2005;85(5):414-419.
27. Henkel V, Mergl R, Kohnen R, Allgaier A, Möller H, Hegerl U. Use of brief depression screening tools in primary care: consideration of heterogeneity in performance in different patient groups. *General Hospital Psychiatry*. 2004;26(3):190-198.
28. Becker S, Al Zaid K, Al Faris E. Screening for Somatization and Depression in Saudi Arabia: A Validation Study of the Phq in Primary Care. *The International Journal of Psychiatry in Medicine*. 2002;32(3):271-283.
29. Gilbody S, Richards D, Barkham M. Diagnosing depression in primary care using self-completed instruments: a UK validation of the PHQ9 and CORE-OM. *British Journal of General Practice*. 2007; 57(541): 650-652.
30. Piazza Marina, Fiestas Fabián. Prevalencia anual de trastornos y uso de servicios de salud mental en el Perú: Resultados del estudio mundial de salud mental, 2005. *Rev. peru. med. exp. salud publica* [Internet]. 2014 Ene [citado 2022 Mayo 08] ; 31(1): 30-38. Disponible en: http://www.scielo.org.pe/scielo.php?script=sci_arttext&pid=S1726-46342014000100005&lng=es.
31. Saito M, Iwata N, Kawakami N, Matsuyama Y, Ono Y, Nakane Y et al. Evaluation of the DSM-IV and ICD-10 criteria for depressive disorders in a community population in Japan using item response theory. *International Journal of Methods in Psychiatric Research*. 2010;19(4):211-222.

32. Matias A, Fonsêca M, Gomes M, Matos M. Indicators of depression in elderly and different screening methods. *Einstein* (São Paulo). 2016;14(1):6-11.
33. Instituto Nacional de Estadística e Informática. Ficha Técnica-Encuesta Demográfica y de Salud Familiar 2019 [Internet]. Lima; 2020 [citado el 20 de mayo de 2022]. Disponible en: https://proyectos.inei.gob.pe/endes/2019/documentos_2019/FICHA_TECNICA_ENDES%202019.pdf

VIII. TABLAS

Tabla 1: Definición operacional de las variables

Variable	Tipo de variable	Definición conceptual	Definición operacional	Indicador
Anhedonia (700A)	Ordinal	Poco interés en hacer las cosas e incapacidad para sentir placer en actividades que anteriormente le generaban placer.	Durante las últimas 2 semanas, ¿qué tan seguido ha tenido pocas ganas o interés en hacer las cosas? (De ser necesario lea: Es decir, no disfruta sus actividades cotidianas)	<ul style="list-style-type: none"> ● Ningún día: 0 ● Varios días:1 ● Más de la mitad de los días:2 ● Casi todos los días:3
Tristeza (700B)	Ordinal	Se ha sentido triste, irritable o sin esperanza	Durante las últimas 2 semanas, ¿qué tan seguido se ha sentido	<ul style="list-style-type: none"> ● Ningún día: 0 ● Varios días:1 ● Más de la mitad

			desanimada(o), deprimida(o), triste o sin esperanza?	de los días:2 ● Casi todos los días:3
Problemas para dormir (700C)	Ordinal	Problemas para dormir o mantener el sueño.	Durante las últimas 2 semanas, ¿qué tan seguido ha tenido problemas para dormir o mantenerse dormida(o), o en dormir demasiado?	● Ningún día: 0 ● Varios días:1 ● Más de la mitad de los días:2 ● Casi todos los días:3
Anergia (700D)	Ordinal	Sentirse cansado o con poca energía.	Durante las últimas 2 semanas, ¿qué tan seguido se ha sentido cansada(o) o tener poca energía sin motivo que lo justifique?	● Ningún día: 0 ● Varios días:1 ● Más de la mitad de los días:2 ● Casi todos los

Cambios en el apetito (700E)	Ordinal	Poco apetito o comer en exceso.	Durante las últimas 2 semanas, ¿qué tan seguido ha sentido poco apetito o comer en exceso?	<p>días:3</p> <ul style="list-style-type: none"> ● Ningún día: 0 ● Varios días:1 ● Más de la mitad de los días:2 ● Casi todos los días:3
Baja autoestima (700I)	Ordinal	Sentirse mal de sí mismo, tener sentimiento de fracaso o abandono propio o de la familia.	Durante las últimas 2 semanas, ¿qué tan seguido ha sentido mal acerca de si misma(o) o sentir que es una(un) fracasada(o) o que se ha fallado a sí misma(o) o a su familia	<ul style="list-style-type: none"> ● Ningún día: 0 ● Varios días:1 ● Más de la mitad de los días:2 ● Casi todos los días:3

<p>Distraibilidad (700F)</p>	<p>Ordinal</p>	<p>Dificultad para concentrarse en diferentes actividades.</p>	<p>Durante las últimas 2 semanas, ¿qué tan seguido ha tenido dificultad para poner atención o concentrarse en las cosas que hace? (De ser necesario: Como leer el periódico, ver televisión, escuchar atentamente la radio o conversar con otras personas)</p>	<ul style="list-style-type: none"> ● Ningún día: 0 ● Varios días:1 ● Más de la mitad de los días:2 ● Casi todos los días:3
<p>Problemas psicomotores (700G)</p>	<p>Ordinal</p>	<p>Moverse o hablar espacio con otras personas, estar inquieto o intranquilo más de lo normal</p>	<p>Durante las últimas 2 semanas, ¿qué tan seguido se ha movido más lento o hablar más lento de lo normal o sentirse más inquieta(o)</p>	<ul style="list-style-type: none"> ● Ningún día: 0 ● Varios días:1 ● Más de la mitad de los días:2 ● Casi todos los días:3

			o intranquila(o) de lo normal?	días:3
Ideación suicida (700H)	Ordinal	Pensamientos de muerte o de lastimarse e alguna manera	Durante las últimas 2 semanas, ¿qué tan seguido ha tenido pensamientos de que sería mejor estar muerta(o) o que quisiera hacerse daño de alguna forma buscando morir?	<ul style="list-style-type: none"> ● Ningún día: 0 ● Varios días:1 ● Más de la mitad de los días:2 ● Casi todos los días:3
Puntaje del PHQ-2	Cuantitativa		700A + 700B	0-6
Puntaje del PHQ-9	Cuantitativa		700A + 700B + 700C + 700D + 700E + 700F + 700G + 700H + 700I	0-27
Diagnóstico de	Dicotómica	Puntaje mayor o igual a	Puntaje mayor o igual a	● Menor de 10:

depresión basado en el PHQ-9 \geq 10		10 obtenido con el instrumento PHQ-9	10 obtenido con el instrumento PHQ-9	Ausente <ul style="list-style-type: none"> • Mayor igual a 10: presente
Diagnóstico de depresión basado en el PHQ-9 \geq 15	Dicotómica	Puntaje mayor o igual a 15 obtenido con el instrumento PHQ-9	Puntaje mayor o igual a 15 obtenido con el instrumento PHQ-9	<ul style="list-style-type: none"> • Menor de 15: Ausente • Mayor igual a 15: presente
Diagnóstico de depresión basado en el PHQ-2	Dicotómica	Puntaje mayor o igual a 3 obtenido con el instrumento PHQ-2	Puntaje mayor o igual a 3 obtenido con el instrumento PHQ-2	<ul style="list-style-type: none"> • Menor de 3: ausente • Mayor igual a 3: presente
Diagnóstico de depresión mayor basado en el DSM-5	Dicotómica	Diagnóstico de depresión según criterios de DSM 5	Cinco (o más) de los siguientes síntomas presentes durante un período no menor de 2	<ul style="list-style-type: none"> • Sí • No

semanas durante la mayor parte del día:

700A=3,
700B=3,700C=3,
700D=3,700E=3,
700F=3, 700G=3,
700H>0, 700I=3)

siempre y cuando uno de los siguientes síntomas este presente:

700A=3 o 700B=3

Diagnóstico de depresión basado en el CIE10

Dicotómica

Diagnóstico de depresión según criterios de CIE 10

2 de estos 3 síntomas durante al menos 2 semanas: 700A=3, 700B=3, 700D=3 además de al menos dos

- Sí
- No

de los siguientes
síntomas: $700F=3$,
 $700I=3$, $700H>0$,
 $700C=3$, $700G=3$,
 $700E=3$

Tabla 2: Concordancia (kappa) entre las diferentes formas de definir depresión (PHQ-2, PHQ-9, DSM-5 y CIE-10) según edad y sexo*

Definiciones de depresión	Total	Sexo		Edad	
		Masculino	Femenino	18 - 64	≥ 65
PHQ-2 x PHQ-9 ≥10	0,61	0,57	0,62	0,59	0,65
PHQ-2 x PHQ-9 ≥15	0,35	0,33	0,36	0,33	0,40
PHQ-2 x DSM-5	0,24	0,22	0,25	0,22	0,28
PHQ-2 x CIE-10	0,30	0,29	0,31	0,28	0,35
PHQ-9 ≥10 x PHQ-9 ≥15	0,55	0,54	0,54	0,53	0,56
PHQ-9 ≥10 x DSM-5	0,37	0,36	0,36	0,35	0,38
PHQ-9 ≥10 x CIE-10	0,45	0,46	0,44	0,43	0,47
PHQ-9 ≥15 x DSM-5	0,73	0,74	0,73	0,73	0,75
PHQ-9 ≥15 x CIE10	0,72	0,73	0,71	0,71	0,73
DSM-5 x CIE10	0,84	0,84	0,84	0,84	0,83

* p<0,0001 en todos los casos

Tabla 3: Prevalencia ponderada según las 5 diferentes definiciones de depresión (PHQ-2, PHQ -9, DSM-5 y CIE-10) según edad y sexo.

Escala	Total (IC 95%)	Sexo (IC 95%)		Edad (IC 95%)	
		Masculino	Femenino	18 - 64	≥ 65
PHQ-2	10,6 (10,1-11,1)	7,2 (6,6-8,0)	13,7 (12,9-14,5)	9,1 (8,57-9,6)	20,6 (18,6-22,8)
PHQ-9 ≥10	6,8 (6,46-7,33)	4,0 (3,56-4,52)	9,54 (8,86-10,2)	5,73 (5,32-6,17)	14,4 (12,8-16,3)
PHQ-9 ≥15	2,6 (2,34-2,87)	1,55 (1,28-1,88)	3,56 (3,16-4,01)	2,1 (1,84-2,34)	5,9 (4,93-7,26)
CIE-10	2,16 (1,92-2,42)	1,32 (1,06-1,63)	2,93 (2,56-3,36)	1,7 (1,51-1,98)	4,9 (4-6,2)
DSM-5	1,6 (1,41-1,83)	0,92 (0,72-1,18)	2,24 (1,92-2,61)	1,3 (1,11-1,51)	3,62 (2,83-4,63)