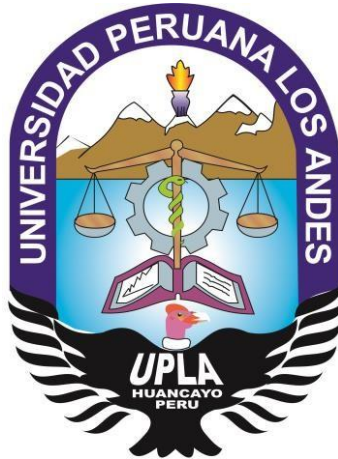


UNIVERSIDAD PERUANA LOS ANDES
FACULTAD DE CIENCIAS DE LA SALUD
ESCUELA PROFESIONAL DE PSICOLOGÍA



TESIS

ESTUDIO DE PRECISIÓN DIAGNÓSTICA PARA EL
INVENTARIO DE DETECCIÓN CLÍNICA DE BECK (BDI-II) EN
ADULTOS, HUANCAYO-2019

Para optar : *Título profesional de Psicólogo*

Autor : Bach. CADENAS HUIZA, Henry Antony

Asesor(a) : Ps. CUADROS ESPINOZA, José Antonio

Línea de Investigación : Salud mental en contexto pluri y multiculturales

Fecha de Inicio/Término : Septiembre del 2018 a abril del 2019

HUANCAYO – PERÚ
2019

Dedicatoria

Para mis padres, ya que gracias a sus oportunas disposiciones comportamentales han dispuesto el contexto necesario para mi formación personal, consolidando de tal modo e indistintamente a mi historia de aprendizaje con cada uno de ellos, las contingencias que me han vuelto un amante de la ciencia, colocándome de tal modo en contacto con una vida significativa movilizada por el amor, la compasión y compromiso por llevar una vida digna de ser recordada.

-El autor.

Agradecimiento

La presente contribución científica ha estado al igual que la ciencia, movilizada por el amor en sus distintas formas de expresión humana, en tal sentido, agradezco infinitamente el soporte académico que mis maestros brindaron a lo largo de mi formación profesional, a mi familia, en los que reside todos aquellos valores que movilizaron mi conducta y especialmente de aquél amor que fue el núcleo de mi inspiración, aquella singularidad espacio-temporal que permitió la constitución de lo que hoy se condensa en el presente trabajo.

-El autor.

INTRODUCCIÓN

La OMS estima que la depresión alcanzará el segundo lugar a nivel global de entre todos los trastornos responsables de la carga de enfermedad para el 2020, siendo considerada como la decimotercera causa más importante que contribuye a los Años de Vida Ajustado en función a la Discapacidad (AVAD) y la primera cuando hablamos de Años de Vida Productiva Perdidos por Discapacidad (AVPD); en nuestro país, más que un problema de salud mental, es un problema de salud pública en riesgo de crecimiento exponencial por la incidencia crítica que reportan diversos sistemas de salud, por lo que viene generando importantes costos para el estado respecto a su diagnóstico, tratamiento y seguimientos eficientes, estos últimos entendidos como procesos consecutivos no aislados, siendo de interés para la presente el componente ligado al diagnóstico, ya que este último representa el eje de la práctica clínica, constituye en tal sentido el principal reto académico en el ámbito clínico de la investigación (Ybarra, et al., 2015).

En este mismo camino y con referencia de la presente, cualquier estudio de precisión diagnóstica es por antonomasia un estudio de *eficacia diagnóstica*, de la cual se desprenderán parámetros matemáticos que otorgarán, en resumidas cuentas, el fundamento científico basado en evidencias que se requiere para la toma de decisiones en la práctica clínica ligada a la puesta en marcha de estrategias de evaluación diagnóstica sean complementarias o centrales a la intervención psicológica vista como una secuencia particular de procesos, determinados por el estudio de caso único (Gonzalo, 2013).

En paralelo y con relevancia para el estudio, las múltiples intervenciones dedicadas a lo largo de esta última década a manejar la variable presentada como el instrumento diagnóstico en el marco de los Tratamientos Psicológicos Basados en Evidencia no han logrado en la

mayoría de sus variantes y específicamente de tradición Conductual, Inter-Conductual, Cognitivo y Contextual (tradiciones fundamentadas en evidencia científica consistente) controlar el sesgo ligado a la precisión en la evaluación diagnóstica y su continua monitorización para el contraste ulterior del impacto terapéutico que tiene el uso de herramientas diagnósticas (grado de efectividad, eficiencia y eficacia) frente a los componentes de otras propuestas psicoterapéuticas, intervenciones farmacológicas o placebo (Pérez, Fernández, Fernández et al. 2003).

Por lo que haría falta realizar revisiones sistemáticas de múltiples estudios en Análisis de Componentes Terapéuticos, con la finalidad de corroborar si se controló o no dicho sesgo ligado a la actividad diagnóstica (como componente terapéutico central o complementario) a través del uso de herramientas dispuestas de parámetros de precisión diagnóstica suficientes para jerarquizarlos como *estándares de oro* en lo que respecta a la detección clínica orientada en la presente a la Depresión. Es así que, para el presente estudio se desplegaron esfuerzos por controlar el grado de incertidumbre relacionado a la actividad diagnóstica complementaria en nuestra especialidad, pretendiendo validar el desempeño discriminativo de instrumentos diagnósticos frecuentemente utilizados en la práctica clínica regional para detectar depresión, a través de la determinación de los mejores parámetros o características operativas de precisión diagnóstica de la nueva prueba (BDI-II) respecto a la prueba de referencia actual o gold standard (EAMD).

Dichos esfuerzos estuvieron sujetos bajo principios de *Investigación Clínica y Epidemiológica*, delimitado por el *método científico*, de carácter *hipotético-deductivo*, tipo *descriptiva* con *diseño de investigación cuantitativo-no experimental*, por ende *observacional-descriptivo*, *transversal con enfoque prospectivo*; habiendo estimado el tamaño muestral (el

cual fue dependiente de la prevalencia registrada de enfermedad; Depresión=22,5%; ajustado por el Intervalo de Confianza al: 95% y Precisión Absoluta admisible para *estudios de precisión diagnóstica* del: 10%) equivalente a 129 personas adultas (casos), sin embargo el cual fue amplificado en el transcurso de la ejecución del mismo por dos motivos puntuales, el primero referido a reducir la Precisión Absoluta al menos al 8%, por lo que su equivalencia muestral correspondió a 201 casos y el segundo motivo referido al control estadístico de las no respuestas (casos inválidos), por lo que el tamaño muestral se constituyó por 210 casos de pacientes que acudieron a consulta en el servicio de Salud Mental (Psicología y Psiquiatría) del Hospital Regional D.C.Q. Daniel Alcides Carrión, los cuales fueron seleccionados de manera no probabilística por casos consecutivos. Los casos fueron orientados por el diagnóstico inicial de los especialistas del servicio para luego exponerse al tratamiento de la variable (instrumento diagnóstico), la cual implicó la administración de la Escala de Depresión de Zung como gold standard y consecutivamente la prueba en estudio, a señalar, el Inventario Clínico de Depresión de A. Beck (BDI-II); cabe resaltar que el instrumento en estudio (BDI-II) utilizado en la presente atravesó un tratamiento estadístico relacionado a la confiabilidad que exhibía previo (prueba piloto) a la aplicación propiamente dicha, cubriendo en esencia 2 dimensiones de este constructo, por un lado la consistencia interna [Alfa de Cronbach: 0,79] y por el otro la concordancia inter-examinadores [Índice Kappa: 0,71], en los que la prueba en estudio se desarrolló de manera aceptable y buena de acuerdo a los intervalos conformados por convenio científico. El análisis estadístico para contrastar la hipótesis fue distribuido en dos secciones, la primera implicó el Análisis de Curvas ROC con la que se halló las características operativas de precisión diagnóstica en los instrumentos y la segunda los Rangos con signos de Wilcoxon con la que se calculó diferencias estadísticamente significativas entre el resultado de la nueva prueba y el gold standard, resultados que se mostrarán en el apartado pertinente.

La presente tesis se encuentra estructurada y dispuesta estructuralmente de la siguiente manera: en el capítulo I se presenta la descripción de la realidad problemática respecto a la variable en estudio, seguidamente se delimita el problema en dimensiones de contexto, temporales y demás, para luego presentar la formulación de problema, justificando consecutivamente las razones que tuvo la presente tesis para ser ejecutada y resaltando los objetivos que se trazaron; en el capítulo II se constituye por la presentación de antecedentes respecto a la presente, además de las bases teóricas que fundamentan la misma y el marco conceptual asociado a este último; para el capítulo III se presenta las hipótesis manejadas por la presente; en el capítulo IV se describe toda la metodología utilizada, la cual incluye: el método de investigación, tipo, nivel, diseño, población y muestra, técnicas e instrumentos de recolección que fueron utilizados, las consecutivas técnicas de procesamiento para el análisis de datos y los aspectos éticos que rigieron la presente tesis.

Finalmente, es preciso señalar que el estudio estuvo en todo momento movilizado bajo estándares de calidad para la producción científica de *estudios de precisión diagnóstica*, habiendo utilizado la lista de comprobación STARD (Standards for Reporting of Diagnostic Accuracy) y QUADAS (Quality Assessment Diagnostic Accuracy Studies), con los que se ofreció los criterios pertinentes para desplegar los procedimientos implicados en la misma (tanto en la etapa de construcción del proyecto como para la presente en la que se expone el informe final), ya sea para el control y rigor de la metodología implicada en la misma, como para sub procesos (prueba piloto, tratamiento de la variable o aplicación del estudio, recolección de datos, procesamiento estadístico, contrastación de hipótesis, etcétera) y su actual reporte final, por lo que se espera haber aportado conocimientos de calidad y dejando abierta nuevas posibilidades metodológicas de investigación clínica en Psicología (González & Velarde 2012).

CONTENIDO

I.	PLANTEAMIENTO DEL PROBLEMA	1
	1.1. Descripción de la Realidad Problemática	1
	1.2. Delimitación del Problema	15
	1.3. Formulación del Problema.....	21
	1.3.1. Problema General.....	21
	1.3.2. Problemas Específicos	22
	1.4. Justificación	22
	1.4.1. Teórica	22
	1.4.2. Social.....	24
	1.4.3. Metodológica	26
	1.5. Objetivos.....	27
	1.5.1. Objetivo General.....	27
	1.5.2. Objetivos Específicos.....	27
II.	MARCO TEÓRICO.....	28
	2.1. Antecedentes (Internacionales y Nacionales)	28
	2.2. Bases Teóricas o Científicas	36
	2.3. Marco Conceptual.....	126
III.	HIPÓTESIS.....	129
	3.1. Hipótesis General.....	129
	3.2. Hipótesis Específicas	130
	3.3. Variables (Definición Conceptual y Operacional).....	131
IV.	METODOLOGÍA.....	133
	4.1. Método de Investigación.....	133
	4.2. Tipo de Investigación.....	136

4.3.	Nivel de Investigación	136
4.4.	Diseño de Investigación.....	137
4.5.	Población y Muestra	138
4.6.	Técnicas e Instrumentos de Recolección de Datos	145
4.7.	Técnicas de Procesamiento y Análisis de Datos.....	146
4.8.	Aspectos Éticos de la Investigación.....	149
V.	RESULTADOS.....	150
5.1.	Descripción de Resultados.....	150
5.2.	Contrastación de Hipótesis	162
VI.	ANÁLISIS Y DISCUSIÓN DE RESULTADOS.....	169
	CONCLUSIONES	178
	RECOMENDACIONES	180
	REFERENCIAS BIBLIOGRÁFICAS.....	181
	ANEXOS	193
	¹ Matriz de Consistencia.....	193
	² Matriz de Operacionalización de Instrumento	196
	³ Instrumentos de Investigación	198
	⁴ Constancia de Aplicación	200
	⁵ Confiabilidad del Instrumento.....	201
	⁶ Data de Procesamiento de Datos	203
	⁷ Consentimiento Informado	209
	⁸ Evidencias de la Aplicación del Instrumento	211

CONTENIDO DE TABLAS

Tabla 1. Costos de manejo clínico de enfermedades mentales prevalentes.....	10
Tabla 2. Niveles de jerarquía respecto al grado de rigor científico	12
Tabla 3. Prevalencia 2017 Hospital RDCQ Daniel Alcides Carrión.....	17
Tabla 4. Prevalencia 2017 Hospital Nacional Ramiro Prialé Prialé	18
Tabla 5. Prevalencia 2018 (enero-julio) Hospital RDCQ Daniel Alcides Carrión.....	19
Tabla 6. Prevalencia anual de trastornos mentales EMSM.....	44
Tabla 7. Incidencia 2018 (julio) Hospital RDCQ “Daniel Alcides Carrión”	49
Tabla 8. Grupos de trastornos del DSM-IV-TR.....	59
Tabla 9. Categorías de la CIE-10.....	59
Tabla 10. Criterios de validez para tratamientos psicológicos	73
Tabla 11. Niveles de evidencia	76
Tabla 12. Criterio de éxito en la terapia.....	77
Tabla 13. Terapias psicológicas apoyadas por evidencia (depresión)	78
Tabla 14. Tablas de análisis de distribución de contingencias 2x2	91
Tabla 15. Distribución de resultados en un estudio de prueba diagnóstica	92
Tabla 16. Modelos de error al determinar la muestra y formular hipótesis.....	95
Tabla 17. Capacidad de eficiencia para valores del IEP	105
Tabla 18. Distribuciones y parámetros descriptivos de una prueba diagnóstica	107
Tabla 19. Valoración de calidad del Índice Kappa	110
Tabla 20. Aplicaciones de las Curvas ROC para pruebas diagnósticas.....	113
Tabla 21. Ventajas e inconvenientes para Curvas ROC no paramétricas.....	117
Tabla 22. Ventajas e inconvenientes para Curvas ROC paramétricas.....	117
Tabla 23. Diferencias entre declaración STARD y QUADAS.....	123
Tabla 24. Clasificación de NE y GR según Sackett.....	124

Tabla 25. Clasificación de NE y GR según Oxford.....	125
Tabla 26. Análisis de precisión y equivalencia muestral (depresión).....	142
Tabla 27. Coeficiente de Consistencia Interna (BDI-II).....	151
Tabla 28. Coeficiente de Concordancia Inter-Examinadores (BDI-II).....	152
Tabla 29. Estadística Descriptiva Edad (+DE)	153
Tabla 30. Estadística Descriptiva Media-Mediana-Moda	153
Tabla 31. Frecuencias de Distribución (Estado Civil/GI/Sexo).....	154
Tabla 32. Parámetros de Distribución Diagnóstica (EAMD/BDI-II)	157
Tabla 33. Parámetros de Distribución Diagnóstica (EA/EAMD).....	158
Tabla 34. Parámetros de Reproducibilidad Diagnóstica (EA/BDI-II).....	161
Tabla 35. Parámetros de Reproducibilidad Diagnóstica (EA/EAMD).....	161
Tabla 36. Área Bajo la Curva (AUC) BDI-II	163
Tabla 37. Comparación de Eficacia Diagnóstica (AUC).....	165
Tabla 38. Media-DE-Varianza (EAMD/BDI-II)	167
Tabla 39. Prueba de Kolmogorov-Smirnov (contraste de normalidad).....	167
Tabla 40. Prueba de Rangos de Wilcoxon.....	168
Tabla 41. Alfa de Cronbach (escala completa y dimensiones).....	201
Tabla 42. Análisis de Fiabilidad BDI-II (por reactivos).....	201

CONTENIDO DE FIGURAS

Figura 1. Modelo cognitivo de la depresión	39
Figura 2. Modelo integrado espectral de la depresión	54
Figura 3. Clínica del síndrome depresivo	57
Figura 4. Árbol de decisión para el estado de ánimo depresivo	66
Figura 5. Cálculo del IC al 95% para proporciones.....	96
Figura 6. Dependencia muestral para el cálculo de valores predictivos	102
Figura 7. Cálculo matemático de Likelihood Ratio (+)	104
Figura 8. Cálculo matemático de Likelihood Ratio (-)	104
Figura 9. Cálculo del Índice Kappa	109
Figura 10. Análisis de operador de receptor	111
Figura 11. Curva ROC	113
Figura 12. Impacto de parámetro diagnóstico sobre decisión terapéutica.....	120
Figura 13. Cálculo del tamaño muestral basado en la Prevalencia (depresión).....	141
Figura 14. Software para el análisis epidemiológico EPIDAT	148
Figura 15. Frecuencia Según Estado Civil.....	154
Figura 16. Frecuencia Según Grado de Instrucción.....	155
Figura 17. Frecuencia Según Sexo.....	155
Figura 18. Parámetros de Validez Diagnóstica (EAMD/BDI-II)	157
Figura 19. Parámetros de Validez Diagnóstica (EA/EAMD)	158
Figura 20. Parámetros de Confiabilidad Diagnóstica (EAMD/BDI-II).....	159
Figura 21. Parámetros de Confiabilidad Diagnóstica (EA/EAMD)	160
Figura 22. Curva ROC Inventario de Depresión de Beck	163
Figura 23. Curva ROC-Comparación de Eficacia Diagnóstica	165

RESUMEN

La investigación se planteó como problema general ¿Demostrará el Inventario clínico de Depresión de Beck parámetros significativos de precisión diagnóstica al distinguir entre pacientes adultos que presenten depresión de los que no lo presenten, al utilizar la Escala de Depresión de Zung como Estándar de oro? En sentido de que el próximo año la Depresión podría alcanzar el segundo lugar a nivel global como uno de los trastornos responsables de la carga de enfermedad y, además del riesgo exponencial para su crecimiento, es en el Perú un problema de salud pública, por lo que genera importantes costos para el estado, siendo su diagnóstico el componente para el que se aplica menor rigor científico a la hora de tomar decisiones en la práctica clínica. Se mantuvo el objetivo de Determinar si los parámetros de precisión diagnóstica (características operativas del receptor) del Inventario de Depresión de Beck (BDI-II) mostrarían significancia discriminativa respecto a la Escala de Depresión de Zung (EAMD) utilizada como Gold Standard al momento de discriminar entre pacientes adultos que presenten o no depresión. Estuvo sujeto a principios de investigación clínica y epidemiológica, se trata por tanto de un estudio con metodología observacional, no experimental, de tipo descriptivo y transversal con enfoque prospectivo; se calculó del tamaño muestral por criterio metodológico sugerido para este tipo de estudios [dependiente de la prevalencia de enfermedad: Depresión=22,5%; IC: 95%; PA: 10% y valor esperado de Sensibilidad y Especificidad: Se=92,0%/Sp=71,0%], estimando 129 casos, estimación incrementada a 210 casos para reducir PA [9%] y por el control estadístico de casos inválidos [muestreo no probabilístico por casos consecutivos a pacientes adultos (<18) que acudieron al servicio de Salud Mental del Hospital Daniel Alcides Carrión]; la prueba piloto incluyó el análisis de confiabilidad del BDI-II, para los que se determinó la consistencia interna y concordancia inter-examinadores [Alfa de Cronbach=0,79/Kappa=0,71]. Se empleó el análisis de Curvas ROC y los Rangos con signos de Wilcoxon para la prueba de hipótesis, obteniendo

resultados en los que el Inventario de Depresión de Beck, demostró mejores parámetros de validez diagnóstica [Se:85.8%-Sp:77.6%], así como mejor confiabilidad diagnóstica [VP+:72.2%/VP-:88.9%/LR+:3.83/LR-:0.18] y reproductibilidad diagnóstica [Kappa:0,61/IC:0,51-0,72], el área bajo la curva del BDI-II fue superior [0,85= p <0,05] al Gold Standard [0,73], corroborando las diferencias de la eficacia diagnóstica con un valor de significación <0,05 [p =0,0005] a través del análisis de Rangos con Signo de Wilcoxon. Las conclusiones a las que se arribó fueron: El BDI-II demostró mejores parámetros de precisión diagnóstica respecto al Gold Standard, apuntando ser un instrumento eficaz por su capacidad discriminativa entre personas con y sin Depresión, lo que lo vuelve también clínicamente relevante en la práctica clínica. Se recomienda en este sentido, desplegar más estudios ligados a determinar la eficacia diagnóstica de los instrumentos que utilizamos con frecuencia en el campo clínico de la psicología, con la intención de aportar evidencia científica y hallar gradualmente la escala de jerarquía de eficacia para la elección de pruebas complementarias.

Palabras clave: Sensibilidad, Especificidad, Depresión, AUC, BDI-II, Discriminación diagnóstica, Eficacia Diagnóstica, Curvas ROC.

ABSTRACT

Did the investigation think about as general problem it will Demonstrate the clinical Inventory of Depression from Beck significant parameters of diagnostic precision when distinguishing among mature patients that present depression from those that don't present it, when using the Scale of Depression of Zung like Standard of gold? In sense that next year the Depression could reach the second place at global level as one of the dysfunctions responsible for the illness load and, besides the exponential risk for its growth, it is in the Peru a problem of public health, for what generates important costs for the state, being its diagnosis the component for which smaller scientific rigor is applied when making decisions in the clinical practice. He/she stayed the objective of Determining if the parameters of diagnostic precision (characteristic operative of the receiver) of the Inventory of Depression of Beck (BDI-II) they would show significance discriminative regarding the Scale of Depression of Zung (EAMD) used as Standard Gold to the moment to discriminate against among mature patients that present or non-depression. It was subject at the beginning of clinical and epidemic investigation, it is a study therefore with observational methodology, not experimental, of descriptive and traverse type with prospective focus; it was calculated of the size muestral by methodological approach suggested for this type of studies [clerk of the illness prevalence: Depression=22,5%; IC: 95%; PA: 10% and prospective value of Sensibility and Specificity: Se=92,0%/Sp=71,0%], estimating 129 cases, estimate increased to 210 cases to reduce PA [9%] and for the statistical control of invalid cases [sampling non probabilistic for serial cases to mature patients (<18) that went to the service of Mental Health of the Hospital Daniel Alcides Carrión]; the test pilot included the analysis of dependability of the BDI-II, for those that it was determined the internal consistency and agreement inter-examiners [Alpha of Cronbach=0,79/Kappa=0,71]. It was used the analysis of Curved ROC and the Ranges with signs of Wilcoxon for the hypothesis test, obtaining results in those that the Inventory of Depression of Beck, demonstrated better parameters of diagnostic

validity [S:85.8%-E:77.6%], as well as better diagnostic dependability [VP+: 72.2%/VP-:88.9%/LR+:3.83/LR-:0.18] and Index of Kappa [Kappa:0,61/IC:0,51-0,72], the area under the curve of the BDI-II was superior [0,85= $p < 0,05$] to the Standard Gold [0,73], corroborating the differences of the diagnostic effectiveness with a significance value $< 0,05$ [$p = 0,0005$] through the analysis of Ranges with Sign of Wilcoxon. The conclusions to those that you arrived were: The BDI-II demonstrated better parameters of diagnostic precision regarding the Standard Gold, aiming to be an effective instrument for their capacity discriminative among people with and without Depression, what also returns clinically him outstanding in the clinical practice. It is recommended in this sense, to deploy more bound studies to determine the diagnostic effectiveness of the instruments that we use frequently in the clinical field of the psychology, with the intention of to contribute scientific evidence and to find the scale of hierarchy of effectiveness gradually for the election of complementary tests.

Key Words: Sensibility, Specificity, Depression, AUC, BDI-II, Discrimination diagnoses, Diagnostic effectiveness, Curved ROC.

CAPÍTULO I

PLANTEAMIENTO DEL PROBLEMA

1.1. Descripción de la Realidad Problemática

La calidad en materiales de investigación, así como del grado de evidencia que estos ofrecen para hoy en día ha ido depreciando circunstancialmente el valor científico en la investigación clínica y con ello reduciendo el avance paradigmático de la práctica clínica actual, por ende, sumando la necesidad imperiosa de los profesionales de la salud para acumular contenidos de información resultantes de investigaciones con conclusiones poco relevantes en el sentido de no incluir valor revolucionario o práctico (que sea ejecutable) en el contexto del conocimiento científico y el ámbito profesional, no permitiendo más que dichos contenidos universalmente aceptados adopten una naturaleza estática a falta de análisis cíclicos en su valor epistemológico actual. Al respecto Tomas Kuhn señalaba: “La ciencia que, para crecer, se apoya firmemente en uno o varios logros del pasado, cuya vigencia es reconocida por dicha comunidad científica como base para ulteriores trabajos, termina siendo acumulativa” (Alvarado, 2005, p. 145).

El criterio anteriormente descrito catalogaría a nuestra ciencia psicológica como “normal”, es decir, una que asiente su avance y progreso paradigmático en una o más realizaciones científicas acumuladas, realizaciones que fundamenten su práctica actual, en ausencia de periodos de crisis, siendo esta inferencia comprobada por la observación y revisión sistemática de investigaciones realizadas en el plano amplio que destina la evaluación psicológica, poniendo énfasis en el área de evaluación clínica diagnóstica, en la que parte de la construcción de pruebas psicológicas y/o adaptación se encuentran hoy sujetas a los principios del modelo psicométrico y sus variadas teorías sobre la medición de variables psicológicas, esta ascensión ha implicado años de investigación y acumulación de conocimiento de gran valor científico que ha permitido los avances que ahora conocemos y utilizamos; desde las primeras formulaciones al respecto, Guilford mencionaba que la evaluación supone comparar alternativas entre los reactivos o ítems de información disponibles, con respecto a criterios dados, por ende, conlleva para el clínico describir y comparar datos, así como ponderarlos y contrastarlos estadísticamente (Ibáñez, 1997, p. 39).

Desde formulaciones constituyentes como las aportadas por Francis Galton con sus aportes para la psicología diferencial y su respetiva evaluación en la variación de rasgos personales o sus expresiones también variadas, por su parte James M. Cattell introduciendo a nuestra ciencia la aplicabilidad experimental para nuestras variables típicas de estudio, con la presunción certera de que “La psicología no puede alcanzar la certeza y exactitud de las ciencias físicas si es que no se funda en el experimento y en la medición” o finalmente sin olvidar la consolidación por parte de Alfred Binet, con la insistencia que mostró por la continua renovación metodológica de nuestra ciencia, así como el interés por los requerimientos prácticos de la ciencia, la constitución del primer test de medida de la inteligencia o la consideración del examen psicológico como complemento de la exploración médica

diagnóstica (Ibáñez, 1997, p. 43) y entre otros grandes aportes que sentaron las bases que hoy guían nuestra práctica profesional en relación al proceso de evaluación en nuestra ciencia.

Al respecto, por lo que hoy conocemos, la evaluación psicológica y sobre todo en la construcción psicológica de pruebas se hallan directrices teóricas aceptadas por nuestra comunidad científica como parte del procedimiento estándar, sin embargo, antes de continuar se hace necesario dar una definición certera de lo que es un test psicológico, al cual podemos de acuerdo a Martínez et al. (2014) entender como “Un instrumento evaluativo o procedimiento en el que se obtiene una muestra de una conducta de los examinados en un dominio especificado y posteriormente es evaluada y puntuada usando un procedimiento estandarizado” (p. 18) fundamentado en procedimientos estadísticos que ratifiquen la veracidad de dichas puntuaciones, es decir, validez y fiabilidad. De este marco conceptual derivan muchos problemas para la interpretación de resultados y limitaciones en el alcance evaluativo, siendo estas sólo algunas: las medidas siempre contienen error puesto que se basan en muestras limitadas que presentan un alto grado de heterogeneidad por tratarse del comportamiento humano, por otro lado, al no contar con una aproximación única a la medición de un constructo, la evaluación puede llevar a distintas conclusiones sobre la caracterización de los sujetos en el atributo o rasgo latente y perder o reducir la posibilidad de generalizar nuestros hallazgos. Muchas de estas problemáticas han sido ya abordadas como se mencionó con antelación por diversas directrices teóricas y estadísticas trascendentes para el manejo de probabilidades, aquí mencionamos algunas que enmarcan al modelo estadístico para la construcción de test:

- Teoría Clásica de los Test (TCT) y sus extensiones, como la teoría de la Generalizabilidad (TG), a veces conocidos como modelos débiles de la puntuación verdadera, basados en el modelo lineal general y en las puntuaciones totales de los test.

- Teoría de la Respuesta al Ítem (TRI) que modeliza las probabilidades de respuesta a un ítem para diferentes niveles del rasgo latente mediante un modelo no lineal.

Posteriormente se ha visto conveniente verificar la calidad de la medición y por ende de los resultados para luego emitir juicios y/o tomar decisiones sobre la persona o grupo de personas a la cual se les realizó la evaluación, siendo esta determinada gracias a la inclusión estadística de las propiedades psicométricas, en las que típicamente y en correspondencia a la teoría psicométrica se pueden hallar según Mislevy: la validez, fiabilidad, comparabilidad, equidad y otros como por ejemplo la determinación de ausencia en sesgos discriminatorios (Martínez, Hernández y Hernández, 2014, p. 23). A partir de aquí podemos analizar parte de la problemática en relación a la presente investigación, puesto que la habilidad de observación y sistematización de hallazgos en materia científica de evaluación clínica se relacionan con un gran y destacado porcentaje de investigaciones centradas a la inclusión de valores psicométricos a través de la evaluación de la validez (p. ej. validez de contenido, de constructo, etc.) y confiabilidad (p. ej. grado de estabilidad, consistencia interna o confiabilidad entre examinadores) de una prueba en construcción o adaptación, enmarcado conceptualmente en términos metodológicos como *Estudio de Propiedades Psicométricas*.

Es así que, la presente investigación no intenta impugnar el cúmulo de conocimiento aportado a lo largo de los avances teóricos y estadísticos que le ha dado la psicometría a la ciencia del comportamiento y análisis de la conducta, a señalar, la Psicología. Sino que pretende amplificar el alcance procedimental, estadístico y finalmente otorgar valores discriminativos de calidad fidedigna del diagnóstico subyacente a contextos clínicos y de la salud que obtenemos como parte del proceso de evaluación diagnóstica en la práctica

profesional cotidiana, incluyendo una metodología de Investigación Clínica y Epidemiológica (fundamentada por su contexto aplicativo), conexas a la Psicometría y además al análisis científico basado en evidencias de los *Estudios de Precisión Diagnóstica*, procedimiento que ofrece al profesional de la salud la posibilidad de añadir criterio y valor científico sobre la veracidad de los resultados diagnósticos que recoge continuamente como parte de su labor estratégica, los cuales guían, amplifican o en algunos casos otorgan soporte a su ulterior juicio clínico respecto a la exploración del paciente, la formulación clínica, la integración terapéutica y la posterior monitorización de la evolución del paciente. En tal sentido, “aplicar *Estudios de Precisión Diagnóstica* a escalas clínicas de evaluación implica, además de brindar consistencia estadística, metodológica y sumar criterio de investigación científica basada en evidencias de segunda jerarquía” (Ybarra et al, 2015, p.4), que las escalas clínicas reciban un tratamiento científico sujeto entre otros al criterio de *falsabilidad*, propuesto por K. Popper como requisito epistemológico para que la actividad de investigar un fenómeno en particular adquiera el carácter de Investigación Científica, la falsabilidad en este sentido, como cualidad indispensable de la actividad científica permitiría que el producto de la presente investigación como enunciado científico se manifieste en concomitancia a la posibilidad de argumentar un juicio clínico producto de la evaluación diagnóstica como verdadero o falso, pudiendo ser confirmado o refutado (Caivano, 1995, p. 6), en analogía a todo procedimiento clínico que pretende, al realizar una exploración para descartar o confirmar la presencia o ausencia de alguna enfermedad en particular, alguna condición sintomatológica, problema de salud o en términos prácticos filtrar a las personas en las que alguna dimensión de su salud se ha visto afectada de las que se encuentran momentáneamente en condiciones saludables.

Ahora bien, se ha incidido hasta el momento en la problemática científica y epistemológica para aclarar el panorama práctico y justificar a su vez la necesidad de aplicar *Estudios de*

Precisión Diagnóstica a escalas psicológicas de evaluación clínica para amplificar la jerarquía de la calidad respecto a su grado de evidencia y aportar consistencia estadística al *Estudio de Propiedades Psicométricas*, sin embargo, es necesario también asociar la problemática anteriormente explorada con la problemática de salud propuesta. Es claro que el espectro de problemas en lo que respecta a la Psicología Clínica y de la Salud es más que diversa y compleja en relación a su exploración clínica, por tratarse de variables de medida acompañados de alto grado de confusión, lo cual no hace más que sumar el interés investigativo de la presente, por lo que para contextualizar el alcance científico y establecer parámetros protocolares de investigación, la presente se enfatiza en esencialmente una entidad sindrómica y las resultantes clínicas comórbidas de la misma denominadas *trastornos mentales y del comportamiento* con su categoría *trastornos del estado de ánimo* (OMS, 2000, p. 18), por lo que nos referimos en cuestión al *síndrome depresivo*, es preciso aclarar antes de continuar que resulta viable para la metodología clínica de la presente investigación asumir el enfoque clínico que permita contemplar al trastorno como variable a estudiar desde una *perspectiva sindrómica* y no nosológica, puesto que entendemos por acuerdo de Llanio (2003) por síndrome a “un grupo de síntomas y signos *no susceptible a ser observada directamente* que se presentan formando un conjunto clínico o “*condición*” que es supuestamente patológica que le da individualidad, pero que *puede obedecer a múltiples causas o etiologías*, a diferencia de la entidad nosológica en la cual el conjunto sintomático se debe a una sola causa” (p. 273), esta perspectiva se ajusta bastante bien a la presente, puesto que los *Estudios de Precisión Diagnóstica* adquieren mayor potencialidad estadística y menor sesgo si no se limitan a establecer parámetros muestrales, entonces, lo que aquí importa es asumir todo el espectro posible de comorbilidades que manifiesten como parte del curso del trastorno, sintomatología depresiva.

Dentro de las múltiples resultantes clínicas para el síndrome, se tiene en principal consideración al trastorno mental y del comportamiento de más impacto en la salud psicológica dentro del sistema de salud y fuera de él, es decir, en la población mundial, por un lado y de manera independiente a la Depresión y por el otro a todas sus variantes más próximas, como entidad representativa de su respectivo conjunto clínico (trastornos del estado de ánimo). Actualmente se presenta antes que como trastorno, como problema de salud pública y de preocupación epidemiológica global, con tasas de incidencia que muestran un crecimiento exponencial respecto a los 8 últimos años hasta la fecha. Los estudios epidemiológicos realizados en los últimos años revelan una alta prevalencia de trastornos mentales en los países de Latinoamérica y el Caribe, además de haber observado un incremento en la carga de enfermedad de los trastornos mentales, en 2002, esa carga ascendió a más del doble, lo cual representó un 22,2%. de dicha carga, la Depresión mayor es considerada como la decimotercera causa más importante que contribuye a los *Años de Vida Ajustados en función de la Discapacidad (AVAD)* y las más importante cuando hablamos de los Años de Vida productiva Perdidos por Discapacidad (AVPD), de ello la OMS ha estimado prospectivamente que para el año 2020, la depresión mayor por si sola alcanzará a nivel global el segundo lugar entre todos los trastornos considerados como responsables de la carga total de enfermedad (Ybarra, et al., 2015, p. 2); de acuerdo con la encuesta mundial de salud mental realizada el 2008, las personas atribuyen mayor discapacidad a los trastornos mentales que a los físicos, debido a la mayor incapacidad social y a las dificultades que éstas generan en su vida (p. 73). En un estudio de Meta-análisis llevado a cabo por Kessler et al (2005), se encontraron que los trastornos del estado de ánimo eran los más prevalentes, alcanzando un 9,5% para la población mundial afectada por algún tipo particular de trastorno del estado del ánimo, (p. 127); los trastornos del estado de ánimo han alcanzado la misma magnitud que los trastornos ansiosos; tienden a la cronicidad y, en muchos casos, promueven la aparición de episodios depresivos y estados

mixtos, además de actuar como factor agravante por su alta predictibilidad para complicaciones posteriores con abuso y dependencia de alcohol y tranquilizantes, además de relacionarse con el desencadenamiento y mantenimiento de múltiples enfermedades médicas.

Según informa el IESM Honorio Delgado (2003) A nivel nacional desde hace 10 años se ha venido incrementando la prevalencia en problemas de salud mental, por lo que el Ministerio de Salud ha comenzado a prestar mayor preocupación por la alta condición de vulnerabilidad que podrían estar manteniendo las poblaciones al interior del país, sin embargo, lo que preocupa en realidad es que en el Perú se realizan pocas investigaciones epidemiológicas con el objetivo de estimar índices de frecuencia mórbida de alcance nacional sobre el estado de la salud mental en la población total, esto muestra además los escasos esfuerzos académicos por realizar estudios clínicos y epidemiológicos de gran alcance, lo cual justificaría en última instancia los insuficientes antecedentes en lo que literatura académica respecta, limitándose netamente a investigaciones locales, focalizados con frecuencia en Lima, existiendo hasta el momento sólo una investigación epidemiológica realizada en la sierra Peruana, la cual comprende las ciudades de Ayacucho, Cajamarca y Huaraz (p. 33). De acuerdo con la información existente, más de un tercio de la población, es decir, casi el 37,3%, tanto en Lima como en ciudades de la sierra, presentó algún trastorno mental registrado por el sistema de salud nacional; así mismo, casi la quinta parte de la población de Lima (19%) y la sexta parte de la población de la sierra (17%) ha sufrido depresión en algún momento de su vida; hasta aquí es posible evidenciar el impacto en la población, puesto que sólo la depresión por cuenta propia asciende a casi la mitad de la población en Lima y la sierra que se ve afectada; de conformidad con el Minsa (2010), “los trastornos mentales y del comportamiento, en particular la depresión, constituye la primera causa de enfermedad en el país, lo cual la vuelve responsable de la pérdida de casi un millón de años de vida saludables” (p. 40).

Respecto a la publicación de reportes epidemiológicos actualizados, divulgados gracias a los diversos medios de comunicación en el país que condensan la información de múltiples entidades asociadas a la salud, de ellas se sabe que en el Perú 8'486,184 personas han sido afectadas por al menos un trastorno mental a lo largo de su vida, de este grupo, para el 2016 se registraron 186,801 casos de depresión atendidos sólo en Lima (Medina, 2016, párr. 2), del total de personas afectadas sólo el 10% acude a los servicios de salud para ser evaluados, de este grupo sólo un 20% recibió un tratamiento adecuado, de la población vulnerable que fue afectada por depresión sólo el 25,8% recibió atención profesional, 21,2% no sabe si necesita ayuda y el 12,7% cree haber necesitado mucha ayuda; este contexto ha generado un incremento mínimo del presupuesto anual por persona para la atención de salud mental en el Perú en tan sólo un 1% respecto a los años anteriores (Perú21, 2015, párr. 5), se asignó 14 soles con 50 céntimos de acuerdo al presupuesto general per cápita establecido para el año 2017, sin embargo, en relación a las regiones la de Moquegua asignó 51 soles por persona siendo la región que más invirtió en salud mental, en comparación de Lima provincias (no incluye Callao: asignó 5 soles con 20 céntimos per cápita), la cual sólo asignó 2 soles con 40 céntimos por persona, la región Junín también mostró precariedad en su inversión, ya que sólo asignó 3 soles con 70 céntimos por persona (Al editor, 2017, 1), estos datos revelan parte de la realidad nacional en lo que salud concierne y sobre todo en lo que a la depresión respecta. Contrariamente, en el 2015 se realizó una investigación respecto a costos para el estado de enfermedades mentales prevalentes, como es de suponer de entre el grupo de enfermedades mentales con tasas de incidencia elevada para aquel año se encontró en segundo lugar a la ansiedad con una incidencia del 19,5% y en tercer lugar a la depresión con una tasa de incidencia del 9%; de dicho estudio se concluyó que el costo total generado por enfermedades mentales prevalentes ascendía a 21'922,517 dólares, ocasionado por el 58% de costos por

depresión (12'621,897 dólares), tales costos para el estado representan la suma de costos específicos en actividades preventivas, diagnóstico/evaluación (depresión: 754,990 dólares), tratamiento y seguimiento (Mosqueira, 2015, p. 813); se muestran en resumen las cifras correspondientes a los costos específicos en la tabla 1.

Tabla 1

Costos de manejo clínico de enfermedades mentales prevalentes (US\$).

Condición Especifica	Población Objetivo	Costo Prevención	Costo Diagnóstico	Costo Tratamiento	Costo Seguimiento	Costo total
Esquizofrenia	4,370	0	239,697	1,653,407	1,458,875	3,351,980
Ansiedad	25,650	943,888	116,424	715,982	263,999	2,040,292
Depresión	24,334	0	754,990	9,697,247	2,169,660	12,621,897
Alcoholismo	21,228	0	665,485	2,966,845	276,018	3,908,348
Total	75,584	943,888	1,776,595	15,033,481	4,168,552	21,922,517
		4.31%	8.10%	68.58%	19.01%	100.0%

Fuente: Mosqueira et al. (2015). Costo de enfermedades mentales prevalentes en Perú (p. 813).

Es probable que una de las tantas posibilidades explicativas para los costos exorbitantes, se justifiquen correlativamente a las altas tasas de infradiagnóstico y sobrediagnóstico en la consulta clínica elaborada por parte de los múltiples profesionales dedicados a la asistencia de la salud mental o al equipo médico subyacente, responsable de tamizajes o procedimientos de descarte previo a la consulta con el especialista, esto último, de acuerdo al nivel de atención y complejidad resolutive de la institución a la cual nos vayamos a referir. Por lo que como refiere Manes et al. (2010) el aumento de diagnósticos registrados en los servicios o departamentos de salud mental pueden hallarse relacionados directamente con el incremento de gastos sanitarios debido al aumento de prescripciones farmacológicas, procesos terapéuticos y con ello el incremento en la frecuencia de utilización de escalas o inventarios clínicos para la evaluación, al respecto diferentes estudios demuestran tanto un sobrediagnóstico en particular para la depresión en atención primaria, como una infravaloración en el campo de la salud, sobre todo en niveles de atención superior e incluso especializados, lo cual a pesar de que el índice de mortalidad para depresión sea bajo, puede ocasionar un considerable impacto socioeconómico secundario a la discapacidad central que genera (p. 48) y, teniendo en consideración que debido

a su carácter patogénico, suele evolucionar con curso crónico incrementando su posibilidad de presentarse en comorbilidad a la ansiedad, por tanto, duplicando el posible riesgo para las no respuestas en evaluación, la poca respuesta al tratamiento, la facilidad para las consecutivas complicaciones e incremento de presentación de cuadros paroxísticos, así como del incremento de recaídas o favorecimiento a la resistencia del conjunto clínico en cuestión. En tal consideración, la atención primaria se sitúa en contextos de la salud como el nivel idóneo para detectar la presencia de sintomatología depresiva, así como los niveles de atención especializada son los más competentes para el diagnóstico diferencial y delimitación clínica para casos únicos, en consecuencia, la elaboración de pruebas de cribado o en su defecto la validación de pruebas diagnósticas puede con mayor certeza, reducir estas tasas en crecimiento de infra y sobrediagnóstico clínico, puesto que en la atención primaria al parecer los especialistas han desarrollado capacidades para detectar casos clínicos (entre ellos: depresión y conjuntos clínicos derivados) con una gran sensibilidad (“mayor” eficacia para detectar a las personas con la enfermedad, condición sintomatológica o problema de salud) a costa de una baja especificidad (“menor” eficacia para excluir correctamente a las personas sin la enfermedad), fenómeno inversamente proporcional en niveles de atención superior (es decir, gran discriminación para especificidad y baja en sensibilidad), lo cual sugiere por consiguiente que el profesional de la salud situado en atención primaria, niveles superiores o incluso especializados, desarrolle habilidades diagnósticas eficaces, eficientes y efectivas (Ybarra, et al. 2015, p. 9), esto último en el marco de: *la Psicología, la Medicina o las Prácticas de Salud Basadas en Evidencias*, aunque también es preciso aclarar que el profesional puede fortalecer dichas habilidades teniendo como vehículo a la práctica científica, siendo tal vez la investigación clínica y epidemiológica la metodología que se ajusta con mayor precisión a su contexto aplicativo, nos referimos pues, al *contexto clínico* y en función del rigor científico para el diseño de estudios observacionales, los *Estudios de Precisión Diagnóstica*

representarían un nivel apropiado de evidencia en la validación de pruebas o procedimientos diagnósticos.

Tabla 2

Niveles de jerarquía respecto al grado de rigor científico.

Jerarquía	Diseños	Sesgo
I	Revisión sistemática y meta-análisis	+ (mínimo sesgo)
I	Estudios clínicos aleatorios	+
II	Estudios observacionales: cohortes y caso-control	++
III	Reporte de series y casos clínicos	+++
IV	Experiencia clínica	++++ (máximo sesgo)

Fuente: Letelier & Moore. (2003). Medicina Basada en Evidencia (p. 941).

Al respecto y en respuesta a la contrariedad planteada, es preciso señalar que con antelación se han realizado pocas investigaciones que intentaron contemplar parte de la problemática formulada por la presente, con un grado limitado respecto a su intencionalidad metodológica y resolutive, es así que p. ej. en un estudio de validación de la escala de Zung para depresión en Colombia se obtuvo valores de discriminación diagnóstica que facilitaron la conclusión de que la prueba mostraba alta sensibilidad (94,7%) y aceptable especificidad (67,0%) limitado a la identificación de episodios depresivos sólo en mujeres universitarias (Campo, et al. 2005, p 57), lo cual sugiere resultados poco consistentes, además de no comparar los valores discriminativos con una prueba de referencia o Gold Standard, fuera de ello también es posible inferir que los resultados no serían de mucha utilidad clínica, ya que las escalas de Zung son tal vez una de las referentes para la identificación de verdaderos positivos y verdaderos negativos en contextos clínicos. En otro estudio Martin, et al. (2012) determinó la utilidad del inventario de Depresión Rasgo-Estado (IDERE) con muestras de 4 países iberoamericanos en contraste a la prueba de referencia CES-D (Escala del Centro de Estudios Epidemiológicos para la Depresión), de la que concluyeron que ambas pruebas tenían buena precisión, excepto por la muestra colombiana que no mostró ningún resultado dentro de los parámetros aceptables (p. 767), de este estudio podemos rescatar el mayor control respecto a su metodología, sin

embargo, al validar clínicamente una prueba, habría sido oportuno validarla en contextos clínicos para obtener valores más precisos en relación a su constructo, tal inconveniente podría a su vez haber sido controlado por Fletcher y Fletcher (2007) si se hubieran incluido casos y controles, con la finalidad de comparar los valores para la discriminación clínica de personas saludables (controles) de personas no saludables (casos), lo cual permitiría discriminar los falsos positivos, es decir, emitir resultados positivos cuando la enfermedad está ausente (p. 43); un buen ejemplo de dicho control para la determinación de las muestras es el estudio realizado por Beltrán et al. (2012) para validar el Inventario de Depresión de Beck, los cuales distribuyeron su muestra en adolescentes parte de la población mexicana clínica (casos) y no clínica (controles), de la que se obtuvieron buenos valores de discriminación diagnóstica tanto en sensibilidad como especificidad (p. 9), pese a ello, tampoco consideraron la utilización de una escala como prueba de referencia, limitándose solo a la elección como prueba de referencia a los criterios diagnósticos de la CIE-10, lo cual podría añadir sesgo en sus resultados, debido a que los criterios diagnósticos de la CIE son utilizados a través de la experiencia clínica del profesional, medida que presenta alta heterogeneidad en función a cada profesional y cada caso. Otra razón por la que se estima que las investigaciones de validación diagnóstica no alcanzaron la utilidad práctica para la cual fueron diseñadas se deba posiblemente a que no formularon adecuadamente el tamaño de la muestra o en su defecto no la supieron hallar de acuerdo a estándares y criterios de investigación en dicho contexto, tal sería el caso del estudio realizado el 2015 en España para el análisis de sensibilidad y especificidad para la Escala de Ansiedad y Depresión Hospitalaria (HAD) en pacientes con fibromialgia, siendo aplicado sólo a 131 pacientes ambulatorios, obteniendo valores que indicaron que la Escala HAD es útil como instrumento de *screening* (Cabrera, et al. 2015, p. 184), resultados que tendrían que ser generalizados sólo al grupo aplicado, ya que para la selección de la muestra en este tipo de investigaciones se debería considerar una significativa a la prevalencia de casos de ansiedad y

depresión de todo el centro de atención u hospital en el cual se efectuó, además claro de que la metodología sugiere incluir a pacientes con diferentes estados clínicos y patológicos de la enfermedad ya que de lo contrario los valores de discriminación diagnóstica perderían consistencia por el sesgo de selección (Argimon y Jiménez, 2013, p. 272), aspecto no controlado por tal estudio.

Por tales motivos, la presente sería una evidencia justificada de ser un estudio *pertinente* a la problemática de salud actual en razón de los beneficios potenciales que puede aportar en contraste a sus bajos costos para su ejecución, sería además *factible o viable* debido a la posibilidad del acceso a instituciones públicas o privadas prestadoras de servicios de salud para poder realizar el estudio, por otra parte, se resalta los posibles aportes metodológicos como parte de la investigación clínica y epidemiológica en ámbitos del estudio de variables del comportamiento humano, así como los aportes inmediatos a los que podría acceder la población diana, es decir, el deseo por generalizar los resultados a poblaciones vulnerables expuestas al riesgo de desarrollar depresión y ansiedad o a las poblaciones afectadas hasta el momento; identificando mejoras para el primer tipo de población respecto a la mejor calidad de exploraciones clínicas a las que se pueden ver expuestos antes de recibir tratamientos o de sencillamente atravesar un proceso de descarte clínico efectivo y por otro lado mejoras para el segundo tipo de población en relación al incremento de la calidad del servicio de salud que hasta el momento venía recibiendo, la mejora para con sus evoluciones y seguimientos conexos a mejores formulaciones clínicas de caso, también basados en la mejor evidencia.

1.2. Delimitación del Problema

Es, así pues, que el problema a estudiar puede ser contextualizado a nivel del departamento de Junín, más específicamente a la provincia de Huancayo, a través de la descripción en primera instancia de la realidad académica, la cual refleja que los esfuerzos académicos para la investigación clínica y epidemiológica en nuestro contexto es escasa para el ámbito profesional de la Medicina Humana e inexistente para el ámbito profesional de la Psicología Clínica y de la salud, a tal punto que como parte de la revisión sistemática de bibliografía realizada por la presente sólo se encontraron 2 estudios peruanos, en la que en uno de los estudios Ampudia et al. (2017) realizó en asociación con universidades extranjeras para poder determinar la “precisión diagnóstica del MMPI-2 en la personalidad delictiva sometiendo su hipótesis al análisis de curvas ROC”, concluyendo en que dicha prueba mostró precisión moderada para discriminar conductas delictivas de las no delictivas (p. 178); lo cual demostraría necesidad metodológica de ampliar la práctica científica orientada a nuevas metodologías puesto que no se encontró al menos un estudio del mismo tipo al realizado el año 2017 que haya tratado algún factor de estudio similar a la presente (es decir, depresión). Es tal vez el segundo estudio encontrado que ofrece mayor equivalencia con la presente investigación, excepto que no discriminan valores de precisión diagnóstica para el inventario de ansiedad de A. Beck, dicho estudio se realizó el año 2014 por parte del equipo médico especialista del Hospital Nacional Cayetano Heredia y de la Facultad de Medicina de la Universidad Peruana Cayetano Heredia, los cuales al validar una versión en español del Inventario de Beck incluyeron valores para la discriminación diagnóstica, siendo corroborados a través de puntos de corte por el análisis de curvas ROC, obteniendo valores de sensibilidad (87,5%) y especificidad (98,21%) altos (Vega, Coronado y Mazzotti, 2014, p. 98), sin embargo, es evidente que estos valores significativamente altos se deban principalmente al tamaño de la muestra (fue aplicado sólo a 136 pacientes hospitalizados en medicina general) y no a las propiedades directas de la prueba,

por lo que lo que se plantea con la presente es además cubrir todas las limitaciones presentadas por los estudios analizados con anterioridad (variables, metodología, prueba de referencia, muestra, etc.) y sobre todo valorar la escala clínica de detección (BDI-II).

Habiendo delimitado la problemática de la presente en el aspecto académico, se hace más preocupante revisar la realidad concerniente a la situación de salud asociada a la depresión en la provincia de Huancayo, *por lo que se realizó un análisis epidemiológico de los 3 conjuntos clínicos más representativos clínicamente y de mayor presentación en la región a modo de vigilancia epidemiológica en el contexto de estudio*. Los reportes de los que se pudo tener acceso evidencian en general que tanto el conjunto clínico que compone a la depresión (episodio depresivo, leve, moderado y grave), ansiedad (ansiedad generalizada, trastorno de pánico y trastornos de ansiedad sin especificación) y cuadros mixtos (trastorno mixto ansioso-depresivo y otros trastornos mixtos de ansiedad), se constituyen como morbilidades de mayor prevalencia al menos entre el grupo de los 10 primeros trastornos más diagnosticadas en 2 de los hospitales más representativos en la región, debido a su ubicación estratégica y capacidad resolutive para cubrir las demandas sanitarias a nivel de toda la región Junín (Cadenas, 2018; Badillo & Cadenas, 2017).

Para una muestra más detallada de la realidad epidemiológica en la región, el jefe de la Oficina de Estadística e Informática del Hospital RDCQ Daniel Alcides Carrión (Hospital Público) facilitó el acceso al reporte de morbilidad del año 2017, en el que se detalla un registro total de ingresos en todo el año de 64,394 casos, siendo el 2,6% y 2,8% la prevalencia de morbilidad reportada sólo por los servicios de Psicología y Psiquiatría respectivamente, de esta proporción el servicio de Psicología reportó 315 casos que conforman el conjunto clínico Depresión, representando por tanto su prevalencia en un 18,3% del total de morbilidades

reportados por dicho servicio (1720 casos); respecto al conjunto clínico Ansiedad el servicio registró 181 casos, demostrando su prevalencia como el 10,5% de la morbilidad general, finalmente para los trastornos Mixtos de Ansiedad y Depresión el servicio reportó 231 casos, equivalentes al 13,4% de la prevalencia en morbilidad, constituyendo a sólo estos 3 conjuntos clínicos como el 42,2% de la prevalencia registrada en el año 2017, lo cual muestra tácitamente el impacto en la salud pública que estas patologías causan anualmente en la región. El servicio de Psiquiatría del mismo hospital demuestra una realidad aproximada, en el que el 26,2% de la prevalencia se constituye por la suma de los 3 conjuntos clínicos estudiados respecto a su registro total de casos (1844 casos), por lo que la carga de morbilidad de ambos servicios en lo que respecta los 3 conjuntos clínicos para el 2017 fue del 34,0% del total de casos para ambos servicios, siendo su prevalencia más de la cuarta parte de la demanda en salud mental (Cadenas H., 2018).

Tabla 3
Prevalencia 2017 Hospital RDCQ “Daniel Alcides Carrión”.

Estimación de PREVALENCIA: 2017		Población		Recuento (Casos)			
1er Conjunto Clínico=Depresión (F32.0/F32.1/F32.2)		N(Total hospital)	TH=64,394				
2do Conjunto Clínico=Ansiedad (F41.0/F41.1/F41.8/F41.9)		N(Psicología)	P=1,720				
3er Conjunto Clínico=Mixto (F41.2/F41.3)		N(Psiquiatría)	P=1,844				
	Conjunto Clínico	Recuento (Casos)	Total	P	P(a*cc)	P(a*PP)	P(ce*TH)
Psicología	Depresión (d)	315 casos	727	P(d)=18,3%	42,2%	34,0%	2,6%
	Ansiedad (a)	181 casos		P(a)=10,5%			
	Mixto (m)	231 casos		P(m)=13,4%			
Psiquiatría	Depresión (d)	115 casos	485	P(d)=6,2%	26,3%		2,8%
	Ansiedad (a)	218 casos		P(a)=11,8%			
	Mixto (m)	152 casos		P(m)=8,2%			
N (P+P)	3,564	N (CC)	1,212	K=100	IC=95%		

P = Prevalencia.
P(a*cc) = Prevalencia acumulada por conjunto clínico.
P(a*PP) = Prevalencia acumulada por conjuntos clínicos de Psicología y Psiquiatría.
P(ce*TH) = Prevalencia por carga de enfermedad por servicio respecto a total de morbilidad del hospital.

Fuente: Cadenas H. (2018). Recurso Personal (Cálculo manual).

En lo que respecta a los registros epidemiológicos del segundo hospital representativo en la región, la Psicóloga Badillo Maldonado Luisa, jefa del servicio asistencial complementario del Hospital Nacional Ramiro Prialé Prialé (Hospital Privado) y por tanto del servicio de

Psicología, permitió el acceso a su registro epidemiológico del año 2017, el cual fue habilitado (sistema de vigilancia epidemiológica) y monitoreado por mi persona en el transcurso de las prácticas pre-profesionales que realicé ese mismo año, del monitoreo realizado se registró 3010 casos sólo en el área de Consultorio Externo y de este total se contabilizó 427 casos respecto al conjunto clínico Depresión, proporción que representa el 15,3% de la prevalencia de casos, un 20,5% respecto al conjunto clínico Ansiedad (572 casos registrados) y un 17,1% de la prevalencia asignada al conjunto clínico Mixto de Ansiedad y Depresión, equivalente a 478 casos registrados, cabe resaltar que no se tuvo en cuenta al resto de la población en la que el servicio de Psicología tiene alcance (p. ej. Psiquiatría, Medicina Interna, Oncología y demás especialidades médicas con las que cuenta el hospital), es así pues que la suma de casos registrados para los 3 conjuntos clínicos representa sólo en Consultorio Externo el 53,1% de la carga de demandas de salud que el servicio tiene que resolver anualmente (Badillo, 2017; Cadenas H., 2018), es preciso aclarar que las proporciones estimadas para el presente año irían en incremento, sin embargo, no es posible presentar cifras consistentes puesto a que el registro del presente año aún no ha culminado.

Tabla 4
Prevalencia 2017 Hospital Nacional “Ramiro Prialé Prialé”.

Estimación de PREVALENCIA: 2017 (Consulta externa)				Población	Recuento (Casos)		
1er Conjunto Clínico=Depresión (F32.0/F32.1/F32.2)				N(Consultorio E.)	CE=3,010		
2do Conjunto Clínico=Ansiedad (F41.0/F41.1/F41.8/F41.9)				N(Total Hospital)	TH=184,542		
3er Conjunto Clínico=Mixto (F41.2/F41.3)							
	Conjunto Clínico	Recuento (Casos)	Total	P	P(α^{*cc})	P(α^{*CE})	P(ce^{*TH})
Psicología	Depresión (d)	427 casos	1,477	P(d)=15,3%	53,1%	49%	1,6%
	Ansiedad (a)	572 casos		P(a)=20,5%			
	Mixto (m)	478 casos		P(m)=17,1%			
N (CE)	3,010	N (CC)		K=100	IC=95%		

P = Prevalencia.
P(α^{*cc}) = Prevalencia acumulada por conjunto clínico.
P(α^{*CE}) = Prevalencia acumulada por conjuntos clínicos de Consultorio Externo.
P(ce^{*TH}) = Prevalencia por carga de enfermedad por servicio respecto a total de morbilidad del hospital.

Fuente: Cadenas H. (2018). Recurso Personal (Cálculo manual).

En lo que va del presente año (enero-julio) el HDAC ha reportado 107 casos para lo que respecta el conjunto clínico depresión representando un 9,6% de la prevalencia de morbilidad total en el servicio de Psicología (1109 casos), en relación al conjunto clínico ansiedad se estimó la prevalencia equivalente al 14,1% (157 casos) y para el conjunto clínico una proporción cercana (14,3%: 159 casos), siendo la prevalencia acumulada (esto es, la suma de los 3 conjuntos clínicos) un 38,1% de la prevalencia del total de casos reportados por el servicio, situación epidemiológica no disperejo a la proporción reportada por el servicio de Psiquiatría, la cual estima un 31,1% de prevalencia para la suma de los 3 conjuntos clínicos (siendo 1591 casos la morbilidad total registrada), representando para cada servicio (Psicología y Psiquiatría) más de la cuarta parte de la morbilidad en sus respectivos reportes y siendo más del 40% de morbilidad registrada en suma por ambos servicios, lo cual devela el crecimiento exponencial respecto a los reportes del 2017 a pesar que tales proporciones sólo constituyen la realidad epidemiológica de enero a julio (Cadenas H., 2018).

Tabla 5

Prevalencia 2018 (enero-julio) Hospital RDCQ “Daniel Alcides Carrión”.

Estimación de PREVALENCIA: 2018 (Enero-Julio)				Población		Recuento (Casos)	
1er Conjunto Clínico=Depresión (F32.0/F32.1/F32.2)				N(Total hospital)	TH=46,316		
2do Conjunto Clínico=Ansiedad (F41.0/F41.1/F41.8/F41.9)				N(Psicología)	P=1,109		
3er Conjunto Clínico=Mixto (F41.2/F41.3)				N(Psiquiatría)	P=1,591		
	Conjunto Clínico	Recuento (Casos)	Total	P	P(a*cc)	P(a*PP)	P(ce*TH)
Psicología	Depresión (d)	107 casos	423	P(d)=9,6%	38,1%	34%	2,3%
	Ansiedad (a)	157 casos		P(a)=14,1%			
	Mixto (m)	159 casos		P(m)=14,3%			
Psiquiatría	Depresión (d)	137 casos	495	P(d)=8,6%	31,1%		3,4%
	Ansiedad (a)	207 casos		P(a)=13%			
	Mixto (m)	151 casos		P(m)=9,4%			
N (P+P)	2,700	N (CC)	918	K=100	IC=95%		

P = Prevalencia.
P(a*cc) = Prevalencia acumulada por conjunto clínico.
P(a*PP) = Prevalencia acumulada por conjuntos clínicos de Psicología y Psiquiatría.
P(ce*TH) = Prevalencia por carga de enfermedad por servicio respecto a total de morbilidad del hospital.

Fuente: Cadenas H. (2018). Recurso Personal (Cálculo manual).

Por tal motivo la descripción del problema subyace también a las altas cifras concernientes a la prevalencia de los 3 conjuntos clínicos en la región, los cuales a su vez sugieren plantear

un estado de emergencia epidemiológica para ambos hospitales al presentarse como un problema de salud en proyección de crecimiento exponencial, implicando, por tanto, mejoras en el sistema de salud que puedan lidiar con dicho crecimiento, siendo posiblemente la validación de procedimientos diagnósticos (en la que se incluyen: entrevistas diagnósticas, procesos de diagnóstico diferencial, evaluación clínicas, tamizajes, entre otros) el paso esencial por el cual podemos replicar estudios epidemiológicos que aporten cifras (respecto a proporción, tasa, razón, prevalencia, incidencia, etc.) con mayor consistencia estadística en relación al descarte o confirmación de la presencia o ausencia de posibles conjuntos clínicos registrando los casos con un mayor potencial discriminativo; ya que el estudio de prevalencia de alguna enfermedad no será significativa si sólo se reduce a la monitorización de casos en los servicios de presentación típica, p. ej. Depresión en Psicología, Hipertensión Arterial en Cardiología, VIH en Infecciosas o Cáncer en Oncología, cuando en realidad el verdadero reto que afronta el profesional consta en elaborar diagnósticos precisos y correctos de patologías infrecuentes en la consulta u hospitalización de un servicio en específico, p. ej. detectar 2 de 100 casos de Cáncer de cuello uterino en pacientes con SIDA en Infecciosas o descartar patologías altamente frecuentes como 35 casos de posible Trastorno de Ansiedad Generalizada en 40 pacientes Oncológicos en su “*primera hospitalización*”; es indiscutible que tales problemas planteados respecto al diagnóstico pueden ser resueltos a través de la experiencia clínica, en lo que muchos profesionales de la salud llaman “*ojo clínico*”, sin embargo, porqué tendría que ser necesario reducir el soporte que brindan las pruebas clínicas como herramientas complementarias, si de lo que se trata es de brindar un abordaje integral y basado en evidencias.

La preocupación científica de la presente recae en consecuencia en el grado de precisión diagnóstica que muchas de las herramientas de evaluación cuentan y, en especial de la escala clínica para detectar depresión debido a su prevalencia en el sistema de salud; además de la

poca difusión para el uso de esta escala en especialidades médicas que podrían tener contacto con casos de nuestra especialidad los cuales sugieran un reto para el médico en lo que respecta a su identificación inmediata antes de su posible derivación; otro punto relevante radica en la actualización de las pruebas de referencia (de primera elección) en nuestra especialidad, ya que muchas de las pruebas hasta el momento utilizadas manejan constructos teóricos ya desfasados o con poca evidencia, valores estadísticos en confiabilidad y validez por actualizar o reactivos no adaptados a contextos actuales, lo cual sugeriría una inefectiva, ineficaz e ineficiente práctica clínica en el marco de la omisión de prácticas en salud basadas en la evidencia y una falta al código deontológico para ejercer alguna profesión en incompetencia a cubrir las demandas de su servicio de calidad para la sociedad.

1.3. Formulación del Problema

En vista de la descripción objetiva respecto a la realidad de la investigación científica aunado a estudios de validación (precisión) de pruebas diagnósticas, el estado epidemiológico en la problemática de salud pública respecto a la variable dependiente del estudio (depresión) y la conjunción en contexto con la Psicología Basada en la Evidencia (PBE), se formula el siguiente problema de investigación científica:

1.3.1. Problema General:

¿Demostrará el Inventario clínico de Depresión de Beck parámetros significativos de precisión diagnóstica al distinguir entre pacientes adultos que presenten depresión de los que no lo presenten, al utilizar la Escala de Depresión de Zung como Estándar de oro?

1.3.2. Problema(s) Específico(s):

¿Describe el inventario clínico de Depresión de Beck parámetros significativos de validez diagnóstica, respecto a la escala de Depresión de Zung utilizada como prueba de referencia?

¿Describe el inventario clínico de Depresión de Beck parámetros significativos de confiabilidad diagnóstica, respecto a la escala de Depresión de Zung utilizada como prueba de referencia?

¿Describe el inventario clínico de Depresión de Beck parámetros significativos de reproductibilidad diagnóstica, respecto a la escala de Depresión de Zung utilizada como prueba de referencia?

1.4. Justificación

1.4.1. Teórica (científica):

En la presente tesis se aplicó fundamentos teóricos del modelo clínico y epidemiológico de la investigación, específicamente, relacionado a Estudios de Precisión Diagnóstica, el cual se orienta a tratar (en un sentido práctico de la investigación clínica) los problemas relativos al proceso de evaluación clínica, diagnóstica, pronóstica (ya que las herramientas complementarias también pueden aportar valores predictivos del estado de los pacientes, además de que sus resultados guían nuestra formulación pronóstica) y el impacto terapéutico que las pruebas diagnósticas generan (en términos probabilísticos pre y postprueba) sobre la decisión de tratar o no al paciente. Aspectos escasamente valorados por los modelos convencionales (limitados al contexto de la evaluación psicológica) al

construir o adaptar herramientas de evaluación clínica, por lo que con la presente se espera también fomentar la exploración de modelos de investigación que aporten mayor rigor científico al lidiar con problemas ligados a la evaluación clínica, fundamentando de tal manera la elección de pruebas complementarias por criterio de aquellas que tras un Estudio de Precisión Diagnóstica demostraron mejor desempeño diagnóstico, traducidos en parámetros estadísticos analizados con alto rigor científico para el control de posibles sesgos, todo ello por un lado en el marco de la Medicina Basada en Evidencias (MBE) y por el otro en la Psicología Clínica Basada en la Evidencia (PCBE) de la cual se desprende la evaluación clínica basada en “pruebas” (Frías y Pascual, 2003, p. 12), haciendo alusión este último término al nivel de evidencia y nivel de calidad de resultados del cual el clínico dispone al momento de realizar algún cribado, exploración inicial en el curso evolutivo de alguna patología, proceso de diagnóstico diferencial, estudio epidemiológico, monitoreo de evolución patogénica o sencillamente para la administración de pruebas a poblaciones en las que hay sospecha de posibles patologías de crecimiento exponencial.

Tal proceso que se encuentra estrechamente ligado a la eficiencia, efectividad y eficacia para elaborar juicios clínicos sobre un paciente o un grupo de ellos, desafío continuo del facultativo para manejar probabilidades e incertidumbre; por lo que, con la presente, se busca promover la evaluación psicológica con rigor científico, que además tiene en consideración el reto que para el profesional implica explorar diversos conjuntos clínicos en su práctica diaria, facilitando de esta manera la buena práctica clínica y adaptación de la ciencia a nuevas necesidades en el campo de la salud (Javier y Crespo, 2012), validando la precisión de los elementos

diagnósticos para que así el profesional pueda ejercer la Psicología Clínica o de la Salud Basada en la Evidencia.

1.4.2. Social (población intra y extra hospitalaria):

Las poblaciones que fueron directamente beneficiadas con los hallazgos de la presente investigación correspondieron, por un lado, a la comunidad académica en la especialidad clínica o de la salud en Psicología y por otro lado tanto a la población intra como extra hospitalaria afectada por ambos conjuntos clínicos en estudio, la primera por hallarse en contacto con el sistema de salud y exponerse a mejoras para la calidad de su atención (la cual incluye abordaje inicial y con ello algún proceso de evaluación) y la segunda porque en algún momento de su vida la alteración de los determinantes de la salud podrán llevarlo(a) a mantener alguna forma de contacto con el sistema de salud y con ello someterse a procesos de intervención clínica que incluyan evaluaciones sujetas a herramientas que demuestren los mejores parámetros de precisión diagnóstica. Ya que la depresión y ansiedad se presentan hoy en día como problemas de salud pública más que como dos conjuntos clínicos de amplio espectro, sus elevadas prevalencias y sus relevantes repercusiones sobre la salud psicológica y física hacen de ambos, los principales motivos de carga de enfermedad para los trastornos mentales a nivel global, con estimaciones predictivas preocupantes; su tendencia a la cronicidad y a la conjunción de presentaciones mórbidas mixtas, demarcan su resistencia al tratamiento y facilitan su proclividad para el incremento del registro de recidivas como parte de su evolución o en su defecto agudizan sus cuadros naturales incrementando la presentación de paroxismos como principales complicaciones, en consecuencia podrían elevar la tasa de mortalidad por suicidios anualmente

respecto a la Depresión y en general podría representar importantes costes para el estado de acuerdo a la asignación de recursos para la salud mental, tanto para su prevención, diagnóstico, como para su tratamiento o para los costes en sus seguimientos. Su diagnóstico implica costes elevados cuando no son realizados con suficiente rigor científico y en lo que respecta a sus procedimientos resultan no ser intrusivos, se realizan con frecuencia mediante la administración de escalas clínicas para la detección los cuales con frecuencia no cuentan con *estudios de validez diagnóstica* que fundamenten su utilización como primera elección.

Es por ello que, si existiera un estudio de precisión diagnóstica que asigne parámetros de discriminación a pruebas clínicas para depresión, las cuales además de no ser costosas o invasivas para el paciente, no sean difíciles de aplicar e interpretar, mejorarían la práctica profesional en contextos clínicos (incluye el área de la salud) del Psicólogo, volviendo a las pruebas aceptables por los pacientes ya que emiten resultados instantáneos lo cual se resume también un menor tiempo invertido por paciente en consulta y sobre todo porque dichos pacientes se someterían a procedimientos de evaluación basados en la mejor evidencia científica para su discriminación diagnóstica, por lo que, si una prueba demuestra mayor ventaja práctica sobre el patrón de oro, sería preciso utilizarla e incluirla como parte de protocolos o guías de práctica clínica que promuevan servicios de calidad en salud (Greenhalgh, 2016, p. 101). Es así que el presente estudio mantiene la intencionalidad de optimizar la exploración diagnóstica integral y, en consecuencia, reducir la tasa de infradiagnósticos o sobrediagnósticos para conjuntos clínicos depresivos dentro del sistema de salud, favoreciendo, en suma, tanto a la población profesional (asistencial) como a la población intra y extra

hospitalaria (en este último caso limitado a contextos de prevención, descarte, cribado o para su futuro abordaje) afectada por ambos conjuntos clínicos.

1.4.3. Metodológica (innovadora):

Se mantuvo la intención de aportar valores discriminativos en cuanto a la precisión diagnóstica de herramientas psicológicas y, subyacentemente contribuir con valores que justifiquen el diseño de algoritmos de intervención clínica en el marco de futuros Estudios de Rendimiento Diagnóstico, en el cual se desarrollan estimaciones de razón para las probabilidades positivas en la toma de decisiones clínicas antes y después de realizar una evaluación (Hulley et al. 2014, p. 173), direccionando la elección de pruebas al momento de abordar a un paciente desde un enfoque basado en evidencias o en su defecto favoreciendo el diseño de contingencias respecto a valores discriminativos poco significativos o inconsistentes que podrían mostrar algunas pruebas diagnósticas.

Por otro lado, se espera que la presente se tome como antecedente fundamental en la región para motivar al desarrollo de ulteriores investigaciones en lo que estudios de precisión diagnóstica para pruebas psicológicas respecta, siendo esta parte del contexto de investigación clínica y epidemiológica descuidada en lo que respecta a la Psicología Clínica o de la Salud, la cual en último término y en medida de los intereses académicos del recién egresado o de profesional en ejercicio clínico, pueda facilitar el abordaje de fenómenos ligados al contexto clínico, sujeto en consecuencia a la metodología de investigación más idónea para tales casos, hablamos pues de la investigación clínica y epidemiológica.

1.5. Objetivos

1.5.1. Objetivo General:

Determinar si el Inventario clínico de Depresión de Beck (BDI-II) demuestra parámetros significativos de precisión diagnóstica que permita discriminar pacientes adultos (≥ 18 a) que presenten depresión de aquellos que no lo presenten, respecto a la Escala de Depresión de Zung (EAMD) utilizada como Estándar de Oro.

1.5.2. Objetivos Específicos:

Estimar parámetros estadísticos de validez diagnóstica que aporta el Inventario de Beck para la detección clínica de Depresión, respecto a la prueba de referencia.

Estimar parámetros estadísticos de confiabilidad diagnóstica que aporta el Inventario de Beck para la detección clínica de Depresión, respecto a la prueba de referencia.

Estimar parámetros estadísticos de reproductibilidad diagnóstica del Inventario de Depresión de Beck, respecto a la prueba de referencia.

CAPÍTULO II

MARCO TEÓRICO

2.1. Antecedentes (Internacionales y Nacionales)

Es preciso señalar que la búsqueda sistemática de antecedentes académicos relacionados con la presente investigación a nivel internacional fue más que fructífera, debido en primera instancia a la cantidad considerable de estudios divulgados por países Latinoamericanos y en última instancia a la relativa consideración por haber puesto en marcha intereses por incluir mayor rigor científico a herramientas de detección clínica en su mayoría para la Depresión, por lo que se citan a continuación los siguientes antecedentes más relevantes tanto para la prueba de referencia (Escala de Depresión de Zung) como para la prueba en estudio (Inventario Clínico de Depresión de Beck) debido sus hallazgos relevantes para la presente, teniendo en cuenta que los demás antecedentes serán citados como parte de las bases teóricas-científicas.

2.1.1. Antecedentes Internacionales:

Prueba de referencia (Gold Standard):

Campo, et al. (2005), en su estudio de “Validación de la escala de Zung para la depresión”, siendo aplicado a una muestra de 110 universitarias de las facultades de medicina y psicología de la UAB las cuales participaron de manera voluntaria, a las que se les administró la escala de Zung para la depresión EZ-D y la entrevista

estructurada SCID-I para episodio depresivo mayor EDM, de la que se encontraron parámetros de precisión diagnóstica caracterizados por sensibilidad de EZ-D para EDM, 94,7% [IC 95%: 90,5-99,7]; especificidad, 67,0% [IC 95%: 56,3-76,3]; valor predictivo positivo, 37,5%; valor predictivo negativo, 98,4%; la prueba kappa de Cohen, 0,39; el área bajo la curva, 0,856, y la correlación test retest fue 0,848 [p <0,01], concluyendo que la EZ-D es una prueba con alta sensibilidad y aceptable especificidad en la detección de EDM en universitarias, “además de poner énfasis en que dicha prueba es frecuentemente utilizada en investigaciones colombianas, pero que ninguna ofrece una validación formal”.

Campo y Díaz (2005), en su estudio realizado en la Universidad Autónoma de Bucaramanga (UAB) de Colombia, se propusieron “Validar una versión abreviada de la escala de Zung para depresión” (de 20 reactivos) a una de menos reactivos, con una muestra relativamente más amplia que su estudio anterior, incluyendo a 138 estudiantes universitarios a los que se les administró tanto la ED-Z para contrastar el diagnóstico establecido con la SCID-I de episodio depresivo mayor EDM, estimando que para la versión abreviada, el alfa de Cronbach fue 0.845, la sensibilidad 0.92%, la especificidad 0.71%, kappa 0.383 y el área bajo de la curva ROC 0.84 para el punto de corte de 22, concluyendo que la ED-Z de 20 reactivos es un instrumento válido al igual que versión abreviada de 10 reactivos, del presente estudio los autores “resaltan la consideración de que los criterios para EDM han cambiado desde la introducción de la ED-Z y que a pesar de ello su uso continua como instrumento de detección”.

Campo, et al. (2006), realizaron un estudio de “Validación de la escala breve de Zung para tamizaje del episodio depresivo mayor en la población general de Bucaramanga” con la finalidad de cubrir limitaciones que tuvieron sus estudios previos, considerando una muestra probabilística mínima de 266 personas adultas residentes en la zona urbana de Bucaramanga, justificada por el criterio establecido por Chmura en 1992 para la determinación de la muestra de estudios médicos de validación epidemiológica (mínima 260), a los que se les aplicó tanto la ES-Z abreviada (de 10 reactivos que mostraron consistencia interna de 0,803) respecto a la entrevista estructurada del DSM para el episodio depresivo mayor SCID-I como prueba de referencia, de la que determinaron una sensibilidad de 95,5%, especificidad de 70,3%, kappa media de Cohen de 0,415 y 0,898 de área bajo la curva receptor-operador, concluyendo que la escala abreviada presenta propiedades psicométricas similares a la escala completa, considerándose como un buen instrumento de detección, “agregan también que ya siendo validada en poblaciones generales deberían también ser estudiado su desempeño en contextos clínicos”.

Prueba en estudio (Nueva prueba):

Sanz, et al. (2005), realizaron un estudio en España en el cual realizaron una “Adaptación española del Inventario para la Depresión de Beck-II (BDI-II)” a través del análisis de sus propiedades psicométricas, trabajando con una muestra de 305 pacientes con diversos diagnósticos psicopatológicos que fueron atendidos entre 2002 y 2004 en la Unidad de Psicología Clínica y de la Salud de la Universidad Complutense de Madrid, a los que con antelación fueron sometidos a una exploración clínica de 3 sesiones para determinar su diagnóstico y durante

dicha exploración se les administró el Inventario de Depresión de Beck en su segunda versión, de la misma se estimó por coeficiente de alfa 0,89 lo que indica un buen nivel de consistencia interna, para su validez convergente mostró correlación positiva y estadísticamente significativa [$p < 0,001$] con otras medidas de depresión auto aplicadas y respecto a su validez discriminativa correlacionó adecuadamente y estadísticamente significativa [$p < 0,01$] con otras escalas, por lo que sus resultados avalan la validez del BDI-II, concluyendo que el BDI es un “instrumento válido de detección y cuantificación de síntomas depresivos en pacientes, si bien su utilidad como herramienta para el diagnóstico diferencial de la depresión es una cuestión pendiente de investigación”.

Beltrán, Freyre y Hernández (2012), realizaron un estudio en México en el que exploraron la “Validez del Inventario de Depresión de Beck”, validándola en una muestra de casos y controles constituida en total por 385 adolescentes, dicha muestra se constituyó por participantes de la comunidad en general los cuales fueron seleccionados por muestreo de bola de nieve (173), pacientes psiquiátricos externos seleccionados intencionalmente con y sin trastorno depresivo (212), basando su diagnóstico paidopsiquiátrico en los criterios CIE-10, se utilizaron además otros instrumentos como CBAS, ERR y BAI para realizar el análisis estadístico, estableciendo un punto de corte de severidad de 10 para el BDI, se encontró una validez discriminante tras comparar su precisión con los criterios CIE-10 capaz de discriminar a la población con y sin trastorno depresivo con una sensibilidad de 0.92% y una especificidad de 0.59%, “siendo considerado por Swets en 1998 como una medida de eficacia moderada para el diagnóstico, incluso

no siendo estos valores producto del mejor punto de corte (con punto de corte 18; Se: 0.84% y Sp:0.82%), los cuales demostraron mayor precisión diagnóstica”.

Sanz y García (2013), en la ciudad de Madrid (España), realizaron un “Estudio de Rendimiento diagnóstico y estructura factorial del Inventario de Depresión de Beck-II (BDI-II)”, aplicando el BDI-II junto al módulo de los trastornos del estado de ánimo de la Estructura Clínica Estructurada para los Trastornos del Eje I del DSM-IV (SCID-I) a 322 pacientes adultos ambulatorios con diversos trastornos psicológicos, considerando como prueba de referencia al SCID-I y el listado de cotejo de síntomas depresivos, de la que se demostró que “el mejor punto de corte para el BDI-II sería de 25, la cual proporciona un 73,5% de sensibilidad y 80,3% de especificidad, lo que supone un índice de eficacia diagnóstica y de concordancia (Kappa) de 78.9 y 0.45, respectivamente, lo cual de acuerdo al área bajo la curva ROC para el BDI-II de .83, con un intervalo de confianza asintótico al 95% de confianza entre .75 y .91, supondría una capacidad discriminadora buena según el criterio de discriminación propuesto por Swets”.

2.1.2. Antecedentes Nacionales:

En lo que respecta a nuestro contexto, la revisión de antecedentes académicos relacionados a la variable del presente estudio o la metodología referida a la *Eficacia de Pruebas Diagnósticas*, muestra que la actividad investigativa es escasa, a pesar de que este tipo de estudios suelen ser común respecto a su única variante, en la que se estudian el impacto de la prueba en la práctica clínica, toma de decisiones y coste-beneficio, encasillada en lo que se conoce como *Estudios de rendimiento diagnóstico* (Anza et al. 2015), de la que no se encontró ninguna

referencia nacional como antecedente, es así pues que a continuación se describen los 2 estudios más análogos a la presente:

Vega, Coronado y Mazzotti (2014), en la ciudad de Lima, los cuales, en su estudio de “Validación de una versión en español del Inventario de Depresión de Beck”, realizado con una muestra por casos consecutivos de 136 pacientes hospitalizados en pabellones de medicina general del Centro Médico Naval, estimaron la capacidad discriminativa del BDI-IA respecto a la Entrevista Clínica Estructurada para el DSM-IV (Structured Clinical Interview for DSM-IV o SCID), siendo la puntuación promedio del BDI significativamente más alta en los pacientes con depresión mayor que en los que no la tenían [26,71 frente a 6,79, $p < 0,001$]. Tomando 18,5 o 19,5 como punto de corte del BDI para el diagnóstico de depresión mayor, la sensibilidad fue 87,5% y la especificidad 98,21%, de la cual los autores concluyeron acertadamente que “la versión en español del BDI posee propiedades psicométricas adecuadas para la detección de depresión limitado a pacientes con características sociodemográficas y clínicas de su estudio”.

Ampudia, Sánchez y Jiménez (2017), gracias a la asociación de universidades extranjeras junto a la Pontificia Universidad Católica del Perú, en su estudio determinaron la “Precisión diagnóstica del MMPI-2 con la personalidad delictiva”, a través del establecimiento de parámetros de discriminación diagnóstica y validándola por puntos de corte como parte del análisis con Curvas ROC, en 1740 reclusos, procesados y/o sentenciados (casos) por diversos delitos y 870 no reclusos (controles), de la que concluyen que el MMPI-2 exhibe resultados significativos para la precisión diagnóstica de delincuentes y no delincuentes, “bajo los criterios

para el intervalo de confianza adecuado para la sensibilidad y especificidad de una prueba propuesto por Swets la cual establece que si el IC intervalo de confianza no incluye el valor de 0.5 la prueba presentará una precisión moderada”.

Como ya se ha descrito, los estudios de validación de pruebas presentados con anticipación, no incluyen rigor científico del cual demanda un Estudio de Precisión Diagnóstica dentro del enfoque clínico y epidemiológico de investigación, por lo que ningún estudio ha incidido su exploración bajo una metodología clínica propiamente dicha, lo cual no ha permitido lograr estimar parámetros de discriminación diagnóstica del Inventario de Depresión (BDI-II) de Beck (siendo la Escala de Zung de Ansiedad y el Inventario de Ansiedad de Beck las pruebas estudiadas que no revelan estudios para determinar sus parámetros de precisión diagnóstica, sólo limitándose a la estimación de sus propiedades psicométricas) respecto de alguna prueba de referencia. Incluso es preciso denotar las dificultades que presentaron para describir al patrón de oro, en el que en algunos estudios ni siquiera se consideró una prueba de referencia como tal (justificando su elección e incluyendo sus parámetros de precisión producto de la investigación), caso equivalente para las pruebas en estudio, de las cuales no se tuvo en consideración sus valores significativos respecto a su precisión diagnóstica.

Por otro lado, dichos estudios demostraron algunos tratamientos para sus variables con poco control metodológico para posibles sesgos, así como muestras poco relevantes sujeta a contextos no clínicos para la generalización de resultados, considerando que la depresión es hoy en día un problema de salud pública, de la cual se dispone reportes epidemiológicos que marcan su alta prevalencia y finalmente no estuvieron sujetos a estándares de calidad para toda su construcción en lo que respecta el estudio de validación de pruebas, hechos que la presente, no dejara de lado tanto como para su formulación metodológica, como para la determinación de la muestra, recolección de datos, procesamiento de los mismos y su posterior interpretación y discusión respecto a la prueba de hipótesis.

2.2. Bases Teóricas o Científicas

DEPRESIÓN:

La depresión se identifica actualmente como la condición psicopatológica más relevante de entre las múltiples patologías estudiadas como parte del amplio espectro concerniente a la ciencia del comportamiento humano, esto es, en primera línea relativa a la Psicología y con ello todo el conjunto de conocimientos científicos desarrollados a su consideración, conocimientos fundamentados en esfuerzos científicos en lo que actividad investigativa respecta, por lo cual, en nuestro medio se han desplegado recientemente medidas de contingencia académicas para el tratamiento de dichas variables en un sentido metodológico, de manera que el incremento de actividades académicas se han tornado proporcionales y acertados al problema de salud pública que hoy en día la depresión se ha convertido, tanto a nivel nacional como global.

Es por ello, que las investigaciones en materia del campo clínico y de la salud respecto a al estudio de dicha variable no han hecho más que incrementar su actividad, publicación y/o divulgación científica en diversos campos de las ciencias de la salud, pese a ello, es probable que se haya descuidado sin intencionalidad alguna la investigación relacionada a su evaluación y diagnóstico con base en estudios de carácter clínico y epidemiológico que aporte el suficiente rigor científico al inicio y sobre todo en transcurso de su formulación, productora por tanto, de la mejor evidencia científica en concordancia a la detección clínica con precisión diagnóstica. Es así que la presente se incluiría al conglomerado de esfuerzos académicos en el estudio diagnóstico de esta variable y más precisamente en la validación diagnóstica de instrumentos clínicos que demuestren mayor capacidad discriminativa en contextos intra-hospitalarios.

Concepción Clínica:

La proliferación de perspectivas teóricas acerca de la depresión la conceptualizan desde numerosas aristas, sin embargo, al delimitarla conceptualmente en el marco de la psicopatología se hace posible reducir ese número de posibilidades explicativas, entendiéndola por ejemplo en tres sentidos esenciales, en primer lugar, como síntomas, las cuales por su naturaleza puede acompañar la manifestación sintomatológica de diversos trastornos mentales, en segundo lugar, como síndromes, agrupando conjuntos clínicos caracterizados por sus elementos particulares, por ejemplo, anhedonia, labilidad emocional, etc. siendo tales conjuntos expresados en amplios espectros mórbidos (sobre todo para la depresión que es dimensional) y en tercer lugar, como enfermedades independientes caracterizadas etiologías distintivas, cursos naturales en su evolución, psicopatología característica, signos patognomónicos e incluso pronósticos cuantitativamente predictivos (Vallejo, 2011, p. 230).

Por otro lado, es posible también conceptualizar la variable desde una perspectiva más clínica asociada operacionalmente al trabajo clínico en contextos hospitalarios, por lo que la depresión sería entendida como un trastorno del estado del ánimo caracterizado por un afecto fundamental depresivo, acompañado de alteraciones cognitivas e incluso somáticas (Heinze y Camacho, 2010, p. 7), el cual puede presentarse en comorbilidad como cuadros mixtos de ansiedad y depresión. Resulta preciso agregar que, como idea general de depresión, la clínica discrimina una variedad considerable de trastornos depresivos, distinguiéndose entre algunos al trastorno depresivo mayor, trastorno distímico, trastorno depresivo recurrente, trastorno depresivo no específico, a los que habría que añadir la depresión atípica (la cual sugiere un reto diagnóstico), la depresión doble, la personalidad depresiva, el trastorno mixto de ansiedad y depresión, la depresión estacional, subsindrómica y la depresión enmascarada (Pérez, Fernández y Fernández, 2003), todos ellos no correspondientes al mismo plano etiológico pero

si a la complejidad diagnóstica que implica sólo la depresión y otros cuadros representativos en la clínica que fundamentan su extensión patológica y sus diversas formas de presentación.

Si bien se han revisado hasta el momento perspectivas conceptuales ligadas al diagnóstico categorial, profusamente divulgados por los manuales de clasificación internacional de enfermedades mentales (CIE-DSM), “se hace necesario conceptualizar a la depresión como un conjunto clínico configurado por signos y síntomas distribuidos espectralmente para diversas morbilidades como parte de su etiopatogenia misma, perteneciente para alguna desviación conductual identificable” (Chinchilla, Correas et al. 2011), por lo que su ordenación en la práctica clínica a posteriori de alguna evaluación diagnóstica categorial complementaría su validez.

Ahora bien, hasta el momento se ha puntualizado algunas conceptualizaciones alcanzadas en su mayoría por enfoques médico psiquiátricos y psicopatológicos descriptivos producto de la observación y categorización sistematizada de signos y síntomas hallados en la práctica clínica habitual, no obstante, la práctica clínica psicológica y específicamente su enfoque terapéutico ha demandado más que mera descripción sistematizada, profunda y avanzada comprensión fenomenológica de diversas presentaciones comportamentales no saludables, traducidas en condiciones caracterizadas por signos y síntomas entendidas a través de las interrelaciones de mecanismos diversos que constituyen el actuar humano; y de entre los enfoques teóricos explicativos menos consistentes científicamente, podemos en el extremo opuesto hallar al enfoque conductual, cognitivo y contextual que aportan modelos comprensivos producto de la investigación y experimentación consistente para conceptualizar a la depresión, así como el modelo cognitivo para la depresión desarrollada por A. Beck, en la que convergen el enfoque clínico y experimental del comportamiento depresivo, el cual se

conceptualiza a través de 3 elementos esenciales (triada cognitiva, esquemas y errores cognitivos) (Beck, Rush et al. 2010) en interacción al mal funcionamiento del procesamiento de la información debido a la activación de creencias nucleares profundas por acontecimientos vitales, los cuales conducen a síntomas fisiológicos, emocionales o conductas disfuncionales; no planteando exclusivamente que tales cogniciones sean la causa de la depresión, sino que las causas pueden ser debidas a numerosos factores (genéticos, bioquímicos, evolutivos, de personalidad, ambientales u otros) o la interacción de los mismos (Ruiz, Díaz y Villalobos, 2012), en tal sentido Beck fundamentó la construcción del Inventario de depresión (BDI) basando su modelo teórico/terapéutico respaldada por evidencia experimental y estudios continuos que la pusieron a prueba para hoy considerarla como la más difundida de entre las 3 líneas científicas de la Psicología.

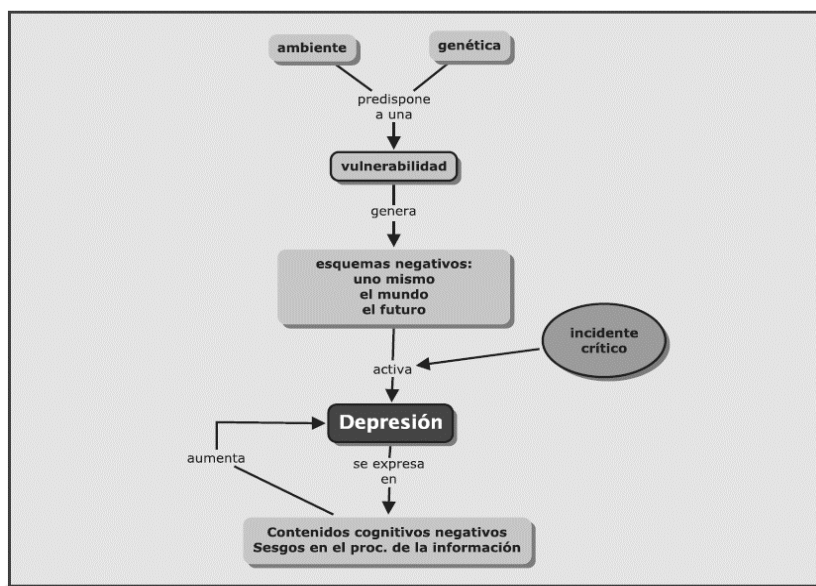


Figura 1 Modelo cognitivo de la depresión.

Fuente: Scott. (2009). Simply effective Cognitive Behaviour Therapy (p. 38).

Recuento Histórico:

Las descripciones acerca de fenómenos subjetivos de sufrimiento, vienen siendo descritas desde hace mucho, se tienen registros tan antiguos como el Antiguo Testamento, en el que se describe un síndrome depresivo sufrido por el rey Saúl, del mismo modo lo demuestra el suicidio de Áyax en la Ilíada de Homero o en diversos relatos mitológicos que reflejan los intentos por describir la melancolía. En promedio al año 400 a. C., Hipócrates utilizó términos como manía y melancolía para describir y referirse a trastornos mentales, como parte de sus estudios enfocados a componentes temperamentales, mencionaba que las personas de la época podían heredar dicha configuración por parte de sus padres. En el año 30 d. C., el médico romano Celso describía la melancolía en su obra *De re medicina*, como una depresión causada por bilis negra. Ya en 1854, Jules Falret describió sujeto a la clínica una afección que denominó *folie circulaire*, en la cual los pacientes experimentaban estados de ánimo de depresión y manía alternantes, continuamente en 1882, bajo contextos de estudio clínicos el psiquiatra alemán Karl Kahlbaum utilizó el término *ciclotimia* para describir la manía y la depresión como estadios de la misma enfermedad (Kaplan y Sadock, 2009), iniciando así lo que sería el enfoque descriptivo en la psiquiatría europea.

Los estudios comenzaron a tornarse prolijos en un sentido favorable de la ciencia y sobre todo de la psiquiatría, por lo que, en 1899, el reconocido psiquiatra Emil Kraepelin describió, partiendo de las teorías de otros psiquiatras franceses y alemanes, la psicosis maníaco-depresiva utilizando la mayoría de los criterios que se usan actualmente para establecer el diagnóstico de trastorno bipolar I, así mismo Kraepelin describió también una depresión que llegó a denominarse como melancolía involutiva y que se ha considerado una forma de trastorno del estado de ánimo que comienza al final de la edad adulta (Kaplan y Sadock, 2009);

Berríos en el año 1988 señaló que el significado que se le dio a la melancolía y manía clásicamente no guardaba estricta relación con su sentido actual:

- Su significado no se establecía en términos polares opuestos.
- Su definición estaba basada en análisis de conducta abierta (aumento o descenso de la motilidad, irritabilidad, etc.).
- Los síntomas no eran definitorios del estado mental.

Por lo que “durante el siglo XIX se produjeron importantes transformaciones de las viejas categorías de manía y melancolía, gracias a los profundos cambios que acontecen en la medicina de la época” (Vallejo, 2011, p. 231); desde el ámbito del conductismo también se han esbozado diferentes hipótesis explicativas de la conducta depresiva, posiblemente, el fenómeno denominado *learned helplessness* o desesperanza aprendida, es uno de los que mejor se ajusta al paradigma conductista de la depresión, Seligman y su equipo de investigación en 1981 descubrieron incidentalmente bajo condiciones experimentales el fenómeno denominado desesperanza aprendida mientras estudiaban el aprendizaje de perros, denotemos que este fenómeno describe un estado caracterizado por la ausencia de conductas adaptativas porque no se reconoce la relación entre unas determinadas respuestas y un alivio de los estímulos aversivos, aunque ha sido poco estudiado experimentalmente en hombres, se ha sugerido que en la depresión humana se recoge una historia existencial caracterizada por un relativo fracaso sistemático en ejercer el control sobre los reforzadores ambientales, lo que lleva a una situación permanente de frustración, siendo entendida la depresión en condiciones en la que el sujeto se percibe a sí mismo como perdiendo todo control sobre tales situaciones, llevándolo a vivenciar inseguridad, pasividad y desesperanza (Vallejo, 2011); por otra parte y más recientemente, diversos esfuerzos en el campo de la psicología conductual con paradigma contextual fueron

aportando nueva evidencia en torno a la depresión, explicándola como un problema de contexto, por interacción de la persona con el ambiente que lo rodea con baja probabilidad de obtener reforzadores positivos o contingencias caracterizadas por relaciones funcionales entre conducta expansivas con refuerzos positivos sociales, situando a la persona en un contexto depresivo (Martell, Dimidjian & Herman, 2013).

Actualmente los estudios están enfocados a establecer marcadores biológicos en muchos trastornos mentales además del estudiado por la presente, aunque no es sencillo mencionarlos pues los pocos estudios encontrados se encuentran en el idioma anglosajón, resultan ser estudios no precisamente orientados a establecer biomarcadores, sino que dicho objetivo es complementario al general, por lo que en muchas ocasiones se encuentran secciones referidas a marcadores biológicos como un campo de investigación por cubrir, motivando continuamente al lector de artículos a poner en marcha estrategias específicas a dicho problema clínico.

Epidemiología:

La distribución de frecuencias asociada a depresión fue descrita en la píldora concerniente al planteamiento del problema, por lo que para este apartado se presentarán registros epidemiológicos más puntuales. Cuando hablamos de distribuciones típicas de morbilidad en el ámbito de la salud, nos referimos principalmente a 2 de las medidas de frecuencia habitualmente utilizadas para el estudio epidemiológico básico, esto es, la prevalencia e incidencia.

La *prevalencia* “es la proporción de individuos de una población que tiene la enfermedad, puede ser estimada a partir de estudios transversales y describe la situación en un momento

determinado, no lo que ocurrirá en el futuro” (Argimon y Jiménez, 2013, p. 328), entonces, es por antonomasia, una proporción del total de la morbilidad registrada en un momento dado.

Los estudios de medida de frecuencias epidemiológicas dependen por tanto del correcto y preciso recuento diagnóstico que realicen los profesionales de la salud en su reporte diario de práctica clínica, la capacidad diagnóstica de procedimientos o herramientas diagnóstica son probablemente el eje problemático al plantear algún estudio de índole clínica, puesto que los reportes indican tasas elevadas de infradiagnóstico en niveles especializados de atención y tasas exorbitantes para sobrediagnóstico en niveles de atención primaria (Manes y Ayuso, 2010), lo cual pondría en duda la fiabilidad de ciertos estudios epidemiológicos realizados, en medida que sus procedimientos o herramientas diagnósticas para la recolección de reportes de morbilidad no incluyen valores de precisión diagnóstica para dar consistencia al registro y discriminación diagnóstica entre un paciente sano y enfermo.

En el caso de la depresión el reporte estadístico habitual es realizado tras la toma de decisiones para el diagnóstico con base en criterios diagnósticos proporcionados por clasificaciones internacionales, ya sea CIE o DSM, por lo que cierta proporción de profesionales de la salud mental la identifican como una prueba de referencia (gold standard) para la elaboración del diagnóstico, dato verificado al revisar estudios de validación diagnóstica o validación discriminativa para ambas variables (Beltrán et al. 2012; Vega et al. 2014; Sanz y García, 2013), “a pesar de que las definiciones de enfermedad van cambiando a medida que cambian las manifestaciones clínicas de diversas patologías” (Greenhalgh, 2016, p. 101) por interacción dinámicas con el ambiente y sujeto. Se ha observado también que las consideraciones de pruebas de referencia o gold standard para estudios epidemiológicos son limitadas respecto a la capacidad diagnóstica que demuestran pruebas como la Escala Zung

Tabla 6
Prevalencia anual de trastornos mentales EMSM.

Trastorno	N.º (%)	EE (IC 95%)
Trastornos de ansiedad		
Trastorno de angustia*	14 (0,3)	0,1 (0,10 - 0,50)
Trastorno. ansiedad generalizada*	8 (0,2)	0,1 (0 - 0,40)
Fobia social *	51 (1,4)	0,1 (1,20 - 1,60)
Fobia específica*	178 (4,6)	0,3 (4,01 - 5,19)
Agorafobia sin trastorno de angustia*	17 (0,5)	0,1 (0,30 - 0,70)
Trastorno de estrés postraumático†	7 (0,2)	0,1 (0 - 0,40)
Trastorno de ansiedad de separación†	31 (1,2)	0,2 (0,81 - 1,59)
Cualquier trastorno de ansiedad†	245 (7,9)	0,5 (6,92 - 8,88)
Trastornos del estado de ánimo		
Trastorno distímico*	14 (0,4)	0,2 (0,01 - 0,79)
Trastorno depresivo mayor*	103 (2,7)	0,2 (2,31 - 3,09)
Trastorno bipolar*	30 (0,8)	0,2 (0,41 - 1,19)
Cualquier trastorno del estado de ánimo	132 (3,5)	0,3 (2,91 - 4,09)
Otros trastornos		
Trastorno negativista desafiante**	5 (0,3)	0,1 (0,10 - 0,50)
Trastorno de conducta**	5 (0,2)	0,1 (0 - 0,40)
Trastorno de déficit de atención**	4 (0,2)	0,1 (0 - 0,40)
Trastorno explosivo intermitente*	69 (1,9)	0,2 (1,51 - 2,29)
Cualquier trastorno del control de impulsos**	65 (3,5)	0,5 (2,52 - 4,48)
Trastornos relacionados a sustancias		
Abuso de alcohol*	49 (1,5)	0,2 (1,11 - 1,89)
Dependencia de alcohol*	15 (0,5)	0,1 (0,30 - 0,70)
Abuso de drogas*	7 (0,2)	0,1 (0 - 0,40)
Dependencia a drogas*	4 (0,1)	0,1 (-0,10 - 0,30)
Cualquier trastorno por sustancias*	57 (1,7)	0,2 (1,31 - 2,09)
Cualquier trastorno		
Cualquier trastorno†	390 (13,5)	0,8 (11,93 - 15,07)
Ningún trastorno†	1411 (86,5)	0,8 (84,93 - 88,07)
1 trastorno†	266 (9,9)	0,7 (8,53 - 11,27)
2 trastornos†	85 (2,6)	0,3 (2,01 - 3,19)
Más de 3 trastornos†	39 (1)	0,2 (0,61 - 1,39)
Severidad		
Severa†	79 (2,3)	0,2 (1,91 - 2,69)
Moderada†	166 (5,6)	0,5 (4,62 - 6,58)
Leve†	145 (5,6)	0,6 (4,42 - 6,78)

Fuente: Piazza & Fiestas. (2005). Prevalencia anual de trastornos y servicios de salud mental (p. 32).

para la depresión, Escala de depresión de Hamilton, etc. para muestras considerablemente grandes, demostrado por estudios de validación que delimitan muestras pequeñas, además de que presenten poca o nula representatividad para generalizar sus resultados a poblaciones intrahospitalarias puesto que la mayoría son validadas en poblaciones estudiantiles o generales (Campo et al. 2005; Campo, Díaz et al. 2005; Campo et al. 2006). Consideración análoga al estudio realizado el 2014 por Montenegro, M. y Kilstein, J. en su estudio de Utilidad del PHQ-9 en el diagnóstico de trastorno depresivo, en la que sólo se incluyeron 152 pacientes, tomando como prueba de referencia al Test de depresión de Hamilton (Montenegro y Kilstein, 2014), la cual no cuenta con suficientes estudios de validación diagnóstica como para ser

considerado patrón de oro para la evaluación de la depresión, estos y demás aspectos serán revisados con mayor diligencia en la pílhora dedicada al diagnóstico.

De entre los estudios más amplios realizados en el Perú respecto a la prevalencia de trastornos mentales, se halla la propuesta intercontinental de estudio ejecutado por Piazza, M. y Fiestas, F. el 2005, como parte de la encuesta anual de trastornos y uso de servicios de salud mental en el Perú, iniciativa de la OMS como parte de la serie de estudios transnacionales y comparativos de 28 países distribuidos en los 5 continentes, específicamente con el fin de recaudar datos empíricos que contribuyan a la toma de decisiones sobre cómo enfrentar de mejor manera la carga pública creciente que implica los trastornos mentales, habiendo ya obtenido la prevalencia anual del 13.9% en México el 2002 y 16% en Colombia el 2003, reportes que significaron mejoras para dichos países, aspecto que no será excepción para Perú, en el que existe poca información acerca de la magnitud de los trastornos mentales, además que la salud mental como aspecto de salud pública no se encuentra adecuadamente priorizada, por lo que la inversión en programas de prevención o promoción se ubica muy por debajo del nivel de inversión de otros países en Latinoamérica, lo cual condiciona perfectamente su inmediata detección; el estudio fue realizado a residentes (de entre 18 a 65 años) habituales en 5 ciudades (Huancayo, Lima, Iquitos, Arequipa y Chiclayo) con un diseño muestral estratificado constituido por 3939 personas, para lo cual emplearon un instrumento que demostró precisión diagnóstica para este tipo de estudios (Entrevista Diagnóstica Internacional Compuesta-CIDI), la cual está estructurada además con base a criterios del DSM-IV y CIE-10 para 17 diagnósticos principales y 6 diagnósticos secundarios, además de incluir módulos que exploran las características sociodemográficas (Piazza y Fiestas, 2014), obteniendo resultados que se resumen en la tabla 6, los cuales exhiben el reporte de la muestra constituida por 3930 personas (a los que se les administró una versión corta del CIDI) y 1801 personas (a los que se les administró la versión larga).

De dicha tabla podemos comprobar que la prevalencia de uno o más trastornos de salud mental en los últimos 12 meses fue de 13,5%, de esta, el 9,9% refiere que presentó al menos un trastorno, el 2,6% de la población al menos 2 trastornos y sólo el 1% menciona haber presentado tres o más trastornos durante ese último año.

Es preciso mencionar que, del reporte total de prevalencia, los tipos de trastornos más frecuentes fueron los trastornos de ansiedad con una prevalencia del 7,9%, seguido por los trastornos del estado de ánimo con un 3,5% de la prevalencia total. Siendo estos dos conjuntos lo más representativos de todo el estudio. Y como era de suponer para el grupo de trastornos del estado de ánimo, el de mayor prevalencia fue el trastorno depresivo mayor en una proporción estimada del 2,7%, seguido del trastorno bipolar con un 0,8%, sin embargo, vale aclarar que el conjunto conformado por cualquier trastorno del estado de ánimo fue prevalente en un 3,5%, lo cual implicaría mayor especificación diagnóstica para poder delimitar adecuadamente dicho porcentaje y asignar sus casos a grupos diferenciados clínicamente, en términos resumidos, sería el grupo conformado por todos aquellos casos que representan un reto diagnóstico, sea por comorbilidad múltiple o atipicidad sintomatológica.

Finalmente, también se detallan indicadores relacionados al grado de severidad, por lo que de entre las personas que presentaron algún trastorno durante ese último año un 41,2% presentaron un cuadro leve, un 41,6% para los casos moderados y 17,2% para los casos severos, de los mismos sólo el 32% de personas con trastornos severos de salud mental recibieron algún tipo de tratamiento en ese último año, respecto a las personas con trastornos moderados de salud mental accedieron a tratamiento sólo en un 18,1%, porcentaje aproximado para el grupo que presentó trastornos leves (15,4%).

En el 2015 se realizaron dos estudios acerca de la prevalencia específicamente de la depresión, esfuerzos distribuidos por la UNMSM y el Instituto Nacional de Salud Mental Honorio Delgado-Hideyo Noguchi; Martina, M. et al. por su parte y Zevallos, E. respectivamente reportaron en distintas poblaciones datos epidemiológicos acerca de la depresión, por un lado, la UNMSM la determinó el año 2015 (estudio publicado el año 2017) en una población adulta en la que identificaron una prevalencia del 14% de depresión como parte de un muestra probabilística de 11,411 casos registrados, identificando a su vez factores de riesgo como: ser mujer con probabilidad del 1,92 para su odds ratio (Odds ratio: Probabilidad de que ocurra un suceso sobre la probabilidad de que no ocurra), ser mayor de 75 años [OR: 1,56], carecer de instrucción [OR: 2,63], vivir sin pareja [OR: 1,32], residir en área rural [OR: 2,12], ser pobre [OR: 2,30] o tener algún discapacidad [OR: 2,63] (Martina, Ara et al. 2017). Por otro lado, Zevallos, E. “estimó a partir de un estudio epidemiológico con una muestra de 3895 adultos pobladores de Sierra urbana y 3031 adultos de Sierra rural que la prevalencia de depresión ascendía al 17,6% y 15,6% respectivamente a los contextos de estudio” (Zevallos, 2015, p. 14).

La *incidencia* “se define como el número de casos nuevos de una enfermedad que se desarrollan en una población de riesgo durante un periodo determinado” (Argimon y Jiménez, 2013, p. 328), por lo cual, la estimación de la incidencia depende en cierta medida de estudios de prevalencia previos o de estudios en los que existe seguimiento de sujetos libres de la enfermedad, permitiendo detectar la aparición de nuevos casos.

Respecto a reportes de la incidencia de depresión es oportuno señalar que son limitadas y, en última instancia deberían ser descritas al contexto más próximo del estudio; hemos descrito hasta el momento que los estudios reportados se centran esencialmente en la prevalencia, pese

a ello, no es imposible determinarla, gracias a los reportes de morbilidad actualizados y facilitados por la oficina de estadística del Hospital RDCQ “Daniel Alcides Carrión” de la ciudad de Huancayo, siendo para este estudio el contexto clínico inmediatamente cercano, el cual da resolución a la principal carga de enfermedad en la región Junín de manera especializada de acuerdo a su nivel de atención, caracterizándose por una población en su mayoría de personas adultas (a pesar de cubrir las necesidades sanitarias de todas las edad; excepto pre y peri natales; de los que se ocupa el hospital Materno Infantil el Carmen situada en la misma ciudad) de estrato socioeconómico bajo (incluso en condiciones de pobreza y extrema pobreza) por tratarse de un hospital público (perteneciente al estado peruano), con la incidencia principal en el grupo de adultos que oscilan de entre los 30 a 59 años correspondientes al ciclo de vida “adulto intermedio” (énfasis añadido), por lo que este representaría en esencia el grupo específico al ciclo de vida en riesgo, situación análoga para los servicios de Psicología y Psiquiatría seguida del grupo de riesgo “adulto joven” (énfasis añadido) los cuales oscilan con edades de entre los 18 a 29 años, de modo que los datos respecto a la incidencia de estos dos servicios se encontrarán bajo la probabilidad de riesgo entre personas adultas que oscilen edades entre los 30 a 59 años, así como adultos jóvenes de entre 18 a 29 años (Cadenas H., 2018).

Es así que las incidencias para ambos servicios (Psicología y Psiquiatría) se estiman con un margen de estudio o seguimiento de 6 meses para hallar la proporción de nuevos casos correspondiente al mes de Julio del presente año en el Hospital RDCQ “Daniel Alcides Carrión” (énfasis de la institución). De este modo el servicio de Psiquiatría se reportó 1410 casos como recuentos totales en el lapso de enero a junio, habiéndose registrado sólo para julio 181 casos, de ellos 15 correspondientes al conjunto clínico depresión, 24 para ansiedad y 18 para el conjunto clínico de presentación mixta, estimando 107, 159 y 115 nuevos casos

respectivamente de los 3 conjuntos clínicos estudiados, hallando por tanto en un 8,3% de incidencia para el mes de julio, respecto a la depresión, 12,9% para ansiedad y 9% en relación a los cuadros mixtos. Proporciones menores respecto a las reportadas por el servicio de Psicología, quienes registraron 983 casos totales como población mórbida o afectada por algún tipo de trastorno mental en el transcurso de enero a junio, siendo 126 casos totales los reportados para julio (depresión=8; ansiedad=19 y mixtos=19), hallando 91, 119 y 121 casos nuevos para depresión, ansiedad y mixtos respectivamente, estimando la incidencia en depresión del 10,2% para el mes de julio, así como un 14% para ansiedad y 14,3% para presentaciones mixtas (Cadenas H., 2018).

Tabla 7
Incidencia 2018 (julio) Hospital RDCQ “Daniel Alcides Carrión”.

Estimación de INCIDENCIA: 2018 (Julio)				Población (Enero/Junio)	Recuento (Casos:)	Casos (Julio)	
1er Conjunto Clínico=Depresión (F32.0/F32.1/F32.2)				N _(Psicología)	P=983	P=126	
2do Conjunto Clínico=Ansiedad (F41.0/F41.1/F41.8/F41.9)				N _(Psiquiatría)	P=1,410	P=181	
3er Conjunto Clínico=Mixto (F41.2/F41.3)							
(Sgto. margen de tiempo equivalente: 6 meses-Persona)							
	Conjunto Clínico	Casos (Enero-Junio)	Casos (Julio)	CN	IA	DI	N(R)
Psicología	Depresión (d)	99 casos	8	91cn	IA(d)=10,2%	40/815 personas-meses.	884c
	Ansiedad (a)	138 casos	19	119cn	IA(a)=14%		845c
	Mixto (m)	140 casos	19	121cn	IA(m)=14,3%		843c
Psiquiatría	Depresión (d)	122 casos	15	107cn	IA(d)=8,3%	46/815 personas-meses.	1288c
	Ansiedad (a)	183 casos	24	159cn	IA(a)=12,9%		1227c
	Mixto (m)	133 casos	18	115cn	IA(m)=9%		1277c
N (P+P)	815	N (cn:Psi)=331 N(cn:Psiq)=381		K=100	IC=95%		

CN = Casos Nuevos. IA = Incidencia Acumulada.
DI = Densidad de Incidencia. N(R) = Población en Riesgo.

Fuente: Cadenas H. (2018). Recurso Personal (Cálculo manual).

Etiopatogenia:

Como sucede con la mayoría de morbilidades relacionadas a la salud mental, sus etiopatogenias han sido poco clarificadas por estudios científicos y, al no contar con evidencia suficiente (acuñada en guías o protocolos de práctica clínica) y consistente en este campo, no se ha hecho más que teorizar posibilidades resumidas en pluricausalidades producto de diversos estudios, pero ninguno específico, es así que “la etiopatogenia de la depresión se ve determinada por múltiples causas, así como diversos factores que dan curso de evolución a dicha entidad, dentro de esta se describen como causas a la genética o causas neuroanatómicas, neurofisiológicas, neuroquímicas, neuroendocrinas, neuroinmunológicas, psicosociales o de personalidad. En primera instancia se da relevancia a la estructura de la personalidad de acuerdo al modelo de vulnerabilidad psicopatológica, finalmente respecto a los factores que permiten la evolución del cuadro se reconocen principalmente a los eventos vitalmente estresantes tempranos, estilos de apego ansioso, ambivalente o desorganizado y aspectos socioculturales” (Cadenas H., 2017, p. 4) que actúan como elementos patógenos dinámicos poniendo en marcha la evolución de la enfermedad.

La dimensión genética de la etiopatogenia de enfermedades mentales está recientemente siendo explorada, por lo que sus aportes podrían ampliar la complejidad de entendimiento, al respecto, David Bartrés y su equipo de investigación determinaron el 2008 los primeros aportes sólidos al campo del estudio de la genética ligada al comportamiento patológico, por lo que aportaron datos genéticos respecto a los trastornos afectivos, en especial para la depresión unipolar por ser la más frecuente, “realizando estudios de familias y de gemelos determinaron que los familiares de los afectados por depresión tienen un riesgo entre 3 y 9 veces superior de padecer algún episodio depresivo, respecto a los estudios de gemelos, se encontró mayor concordancia entre gemelos monocigóticos que entre los dicigóticos” (Bartrés, Redolar et al.

2008, p. 346), en un meta-análisis por ejemplo se explicó como conclusión que el riesgo de padecer depresión unipolar podría ser explicado mayoritariamente por la varianza ambiental específica o no compartida del 58% al 67% y por la varianza genética aditiva del 31% al 42%, con muy poca influencia de la varianza ambiental compartida (del 0% al 5%); otro dato relevante que nos aporta va respecto a la prevalencia según sexos, la cual es desigual puesto que las mujeres mantienen un riesgo mayor de padecer depresión, sin embargo, la heredabilidad no cambia según el sexo, siendo relevante al considerar que los genes que puedan mediar en su etiología son los mismos para ambos sexos.

Debemos, sin embargo, comprender que el estudio de etiopatogenias conlleva mucho material enmarcado en el modelo biomédico, por lo que sus estudios suelen centrarse en tales cuestiones, omitiendo en muchos casos elemento externo o el aprendizaje para su constitución patogénica.

Etiología:

De acuerdo a la naturaleza mórbida del conjunto clínico estudiado por la presente se hace posible inferir que sus configuraciones etiológicas no sólo radican en causas patógenas anatómicas estructurales, desbalances neuroquímicos, fisiopatología neuroendocrina, la configuración genética heredada en el sentido de tendencia y predisposición para desarrollar la enfermedad o en general a todo lo que conforma al nivel de organización biológica, sino que también se disponen en otros niveles de organización, las que corresponden a la psicológica y social, constituyendo de esta manera su carácter pluricausal y por ende, plurifactorial, haciendo del mismo, el conjunto clínico más complejo en su exploración etiológica, a continuación se presentará algunos de los componentes más relevantes en la investigación que configuran estos niveles de organización.

Bioquímicos cerebrales, esta teoría postula que los cambios en las monoaminas cerebrales, entre ellas, la noradrenalina, dopamina y serotonina, condicionan la sintomatología expresada por la depresión, específicamente por un déficit absoluto o relativo de norepinefrina en sinapsis específicas cerebrales, así mismo un déficit de serotonina como prerrequisito etiológico básico de la depresión, cabe resaltar que este enfoque teórico en su mayoría se encuentra fundamentado en estudios psicofarmacológicos (Vélez, Rojas et al. 2010).

Neuroendocrinos, este enfoque se ve respaldada por múltiples estudios que han mostrado anomalías en el eje hipotalámico pituitaria adrenal (HPA), debido a la existencia del aumento del factor liberador de la corticotrofina (CRF) en plasma y en LCR, curva plana de producción de hormona adrenocorticotrópica al estimular con CRF, aumento de cortisol y no respuesta a la prueba de supresión de cortisol con la dexametasona, en diversos pacientes depresivos, un hallazgo muy importante y que persiste una vez mejora la depresión fue el de la comprobación en la disminución de la secreción de la hormona del crecimiento durante el sueño (Vélez, Rojas et al. 2010; Hales, Yudofsky et al. 1996); siendo esta la evidencia con mayores estudio de respaldo, ya que respecto a los demás trastornos específicos no se cuenta con estudios actualizados o con resultados dignos de publicación.

Genéticos y Familiares, se cuenta con estudios familiares y de pedigrí a favor de la existencia de factores genéticos en depresión, por lo que es 1,5 a 3 veces más común en familiares biológicos de primer grado de personas que sufren éste, que la población general, respecto a gemelos monocigóticos es de 37% en hombre y 31% en mujeres, a diferencia de los dicigóticos con 23% y 25% de frecuencia respectivamente, por lo que la heredabilidad en diferentes estudios se calcula entre 36% a 75%, demostrando el significativo componente genético (Vélez, Rojas et al. 2010), demostrando mayor tasa de concordancia entre gemelos

monocigóticos, es así que su naturaleza hereditaria de los trastornos mayores del estado de ánimo sugeriría la necesidad clínica del asesoramiento genético (Hales, Yudofsky et al. 1996) para el control del riesgo a exposición que facilite su desencadenamiento.

Psicológicos y Sociales, muchas de las causas en incremento de incidencia o prevalencia de depresión se encuentra ligado a circunstancias vitales y estrés ambiental, como acontecimientos relacionados a la pérdida de algún familiar de edad temprana o no coherente al ciclo de vida (fallecimiento en la infancia, adolescencia o juventud), también abusos o violencia sistemática que además de depresión genera hiperactivación en estado de alerta por temor al agresor, pérdida del conyugue, desempleo, ausencia de soporte social (Cardelle, Lorenzo et al. 2009), además de la exposición a desastres naturales, guerras, conflictos armados, permanencia en campos de concentración y períodos de desintegración política y social (Vélez, Rojas et al. 2010).

En lo que respecta los factores etiológicos ligados a causas psicológicas es evidente gracias a los estudios basados en evidencia y de significativo rigor científico que las corrientes conductuales, cognitivas y contextuales del aprendizaje aportan la mejor explicación causal, en este sentido, rasgos de personalidad actúan también como rasgos de vulnerabilidad para diversos trastornos, la teoría de aprendizaje atribuye importancia etiológica a trastornos del comportamiento y de la cognición, Seligman propuso por ejemplo que la desesperanza aprendida actúa como elemento central de algunas formas de depresión, Lewinsohn observó que la pérdida de refuerzo en el ambiente puede crear y mantener un comportamiento depresivo, para A. Beck el problema psicológico básico de la depresión es la traducción de un pensamiento distorsionado que provoca una visión o inferencia negativa de sí mismo, de los demás y del futuro, es así pues que las teorías cognitivas relacionan la depresión con

distorsiones cognitivas y las teorías conductuales por mecanismos de indefensión aprendida, acontecimientos externos en el sentido de la baja probabilidad de condicionar un contexto que disponga de refuerzos que configuren un comportamiento depresivo (Vélez, Rojas et al. 2010; Cardelle, Lorenzo et al. 2009).

Actualmente se acepta un modelo integrado en el cual se considera a la depresión desde un punto de vista sindrómico, en el cual las depresiones están situadas en un espectro, con implicaciones etiológicas y tratamientos diferentes (ver figura 2).

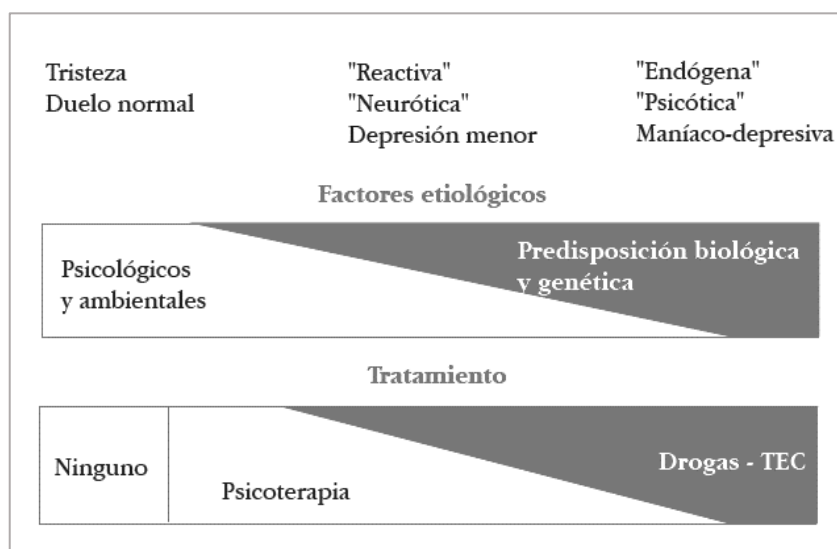


Figura 2 Modelo integrado espectral de la depresión.

Fuente: Vélez et al. (2010). Fundamentos de medicina Psiquiatría (p. 173).

Como podemos constatar, las teorías conductuales, cognitivas o las relativamente recientes contextuales ofrecen un marco teórico, experimental, práctico y científico más amplio y consistente en relación a sus fundamentos, por lo que se infiere una mejor posibilidad de evaluación bajo pruebas, escalas, inventarios, etc. enmarcadas bajo estos principios explicativos científicos actualizados.

Curso Clínico:

En relación al curso de la depresión se estima que la edad promedio de inicio es la tercera década de la vida, pero puede presentarse a cualquier otra edad; se considera uno de los conjuntos clínicos más recurrentes, por lo que su presencia como comorbilidad de otros trastornos también es alto, la manifestación sintomatológica puede remitir por completo o pueden dar a un cuadro crónico, pese a ello, la depresión suele tener una duración larga (Tajima, Molina et al. 2014), su historia natural incluye sintomatología mínima de 2 semanas y máxima aproximada de 6 meses a 14 meses aunque persiste más de dos años hasta en 20% de los pacientes, su finalización puede ser gradual o abrupto, si se prolonga consideramos distimia bajo el criterio relativo de un año para niños y dos para adultos, en general sus antecedentes naturales ha sido ya muy bien descritos, por lo que la depresión se produce de manera episódica en la mayoría de los casos y sus síntomas desaparecen y el estado de ánimo y el funcionamiento se hacen normales o caso normales entre episodios, del 15% al 20% de todos los pacientes presentan cronicidad, por lo que sin tratamiento su depresión no sería resuelta, relacionándose de manera cercana con los intentos de suicidio y suicidio consumado (Elkin, 2010; Dupont, 2005).

Psicopatología y Clínica:

Para explicar la psicopatología de la depresión, debemos antes comprender que el ser humano está dotado biológicamente de estructuras anatómicas que propician el desarrollo de dos afectos fundamentales, a señalar, de tipo depresivo o de tipo expansivo (Hernández, 2013, p. 73), en lo que respecta al afecto depresivo asociado en su estado patológico al trastorno depresivo mayor, expresa la insatisfacción y la tensión de las pulsiones. Su tonalidad cualitativa es la del dolor emocional, la tristeza, angustia y sufrimiento subjetivo. Los afectos depresivos en su estado de regresión más profunda descargan una tensión insoportable que se externaliza

en forma de sentimientos vitales de disgusto, abatimiento e inseguridad, o de emociones paroxísticas de ira, rabia, desespero, miedo o cólera.

En términos cuantitativos, los cuales permiten describir la naturaleza de la sintomatología ya sea en niños, adolescentes o adultos, se puede caracterizar por una respuesta depresiva en ausencia de una causa objetiva aparente que la produzca, en segundo lugar, dicha respuesta es desproporcionada a dicha causa y finalmente el tiempo de presentación habiendo causa es prolongada en el tiempo (Hernández, 2013). Secundariamente a ello se puede desarrollar la siguiente psicopatología: tristeza patológica, apatía, anhedonia, irritabilidad, sentimiento de culpa, angustia patológica, incontinencia afectiva, labilidad emocional, ambivalencia afectiva, síndrome amotivacional, indiferencia afectiva y fatiga.

La *clínica* referida a los trastornos del estado de ánimo se caracteriza en general por la presentación de síntomas anímicos como tristeza: estado de ánimo deprimido la mayor parte del día, sentimientos de vacío, desmotivación, desesperanza o bien síntomas objetivos de tristeza como llanto frecuente, llegando incluso en casos graves manifestarse la tristeza como una anestesia afectiva, con diferencia hacia el entorno e incapacidad para implicarse emocionalmente en nada. Irritabilidad, la cual es más frecuente en niños y adolescentes acompañados de reactividad y actitud opositora. Anhedonia como pérdida o disminución de la capacidad de experimentar placer por actividades realizadas, esta incluye la disminución de la libido y del interés sexual. Respecto a los síntomas somáticos el paciente puede manifestar pérdida acusada de peso y apetito, alteraciones en la consolidación, mantenimiento o conciliación del sueño, así mismo fatiga y pérdida de energía diaria, posiblemente acompañado de artromialgia y neurastenia; en relación de la sintomatología ligada al aspecto cognitivo, se caracteriza por una visión negativista de sí mismo, de los demás y del futuro, lo cual facilita

consecutivamente la formulación de ideación suicida, aunado al de inutilidad o de culpa excesiva, finalmente, los síntomas conductuales se caracterizan por alteraciones en la concentración, atención y memoria, en casos graves inhibición psicomotora e incluso catatonía.

TIPO DE SÍNTOMA	ALTERACIÓN
Síntomas anímicos	<ul style="list-style-type: none"> • Tristeza • Irritabilidad • Anhedonia • Apatía
Síntomas somáticos	<ul style="list-style-type: none"> • Peso y apetito • Sueño • Pérdida de energía • Quejas somáticas • Sexuales
Síntomas del pensamiento	<ul style="list-style-type: none"> • Visión negativista • Ideas de inutilidad y culpa • Depresión delirante (presencia de delirios y alucinaciones)
Síntomas conductuales	<ul style="list-style-type: none"> • Concentración y atención • Memoria • Psicomotricidad • Descuido del aseo/cuidado personal

Figura 3 Clínica del síndrome depresivo.

Fuente: Tajima et al. (2014). Manual de Psiquiatría (p. 818).

Diagnóstico y Propedéutica:

Gran parte de la proporción en la elaboración diagnóstica para la depresión como conjunto clínico o en general cualquier trastorno mental, se establecen a través de una serie de formulaciones clínicas descriptivas (de carácter categorial o dimensional), por lo que las pruebas analíticas de laboratorio, estudios por imágenes, identificación de marcadores biológicos o cualquier otro cúmulo de procedimientos médico-clínicos (de carácter tecnológico avanzado) para el diagnóstico patológico no se ha logrado extrapolar con éxito hasta el momento al hablar del diagnóstico sobre el comportamiento humano.

En este sentido, existen múltiples series de síntomas, signos, síndromes e incluso signos patognomónicos a estadios de la enfermedad, como el acompañamiento de la fiebre al síndrome

confusional, el cual equivale clínicamente al Delirium (Llanio et al. 2003; Hernández, 2013), los cuales son producto del registro sistemático a través de la exploración y descripción en la historia de la práctica clínica. Por ende, al tratarse de conjuntos clínicos sensibles a la descripción del profesional, su determinación diagnóstica mantiene la posibilidad de formularse por medio de múltiples técnicas, procedimientos o instrumentos, por lo que el proceso diagnóstico puede ir de un análisis longitudinal retrospectivo de la historia clínica hasta las muy bien conocidas entrevistas diagnósticas (Vásquez y Muñoz, 2002), las cuales han aportado sistematización al proceso de exploración más antiguo en la práctica clínica producto de la investigación. Ese esfuerzo por la sistematización de la semiología en nuestra especialidad ganó mayor consistencia científica al incluir sistemas de clasificación internacionales como categorización de criterios clínicos para el diagnóstico. Siendo su objetivo el de facilitar la comunicación interdisciplinar, constituir una base de conocimientos para obtener y compartir información relacionada a la salud mental, aportar conocimientos descriptivos sobre las entidades mórbidas en cada especialidad, realizar predicciones y formular teorías (Vallejo, 2011), cumpliendo de esta manera con los 3 propósitos relevantes de poder contar con un sistema de clasificación, a saber, lenguaje científico, determinación terapéutica y comprensión de entidades mórbidas (Hales, Yudofsky & Talbott, 1996).

Hoy en día los dos sistemas diagnósticos más importantes son la Clasificación Internacional de Enfermedades en su décima versión (CIE-10), desarrollada por la OMS y el Manual Diagnóstico y Estadístico de Enfermedades Mentales en su 4ª edición de texto revisado promovido por la APA, ambos con su sistema de clasificación multiaxial y no axial, permitiendo codificar la condición de salud y afección respecto a la distribución de ejes clínicos; ambos manuales son a la fecha utilizados por profesionales de la salud mental (en general profesionales en ciencias de la salud) como medio para guiar su formulación clínica

para el diagnóstico, sin embargo, actualmente ya se encuentran habilitadas nuevas versiones, respecto al DSM su quinta versión se publicó oficialmente el 2014, respecto a la CIE se planificó la publicación de su onceava versión para enero del presente año, sin embargo, al parecer se extenderá hasta mediados del 2019, este incluiría un nuevo sistema diagnóstico, en el que se vislumbra a las entidades mórbidas desde el modelo dimensional o espectral de la psicopatología, modelo que comienza a mostrar tendencia teórica y práctica en algunos estudios psicopatológicos que demuestran mayor fiabilidad y validez diagnóstica al modelo categorial, considerado desde hace muchos años como el convencional (Hernández, Palacio y Freyre, 2011, p. 115). La distribución de trastornos por categorías o grupos mórbidos se presenta de la siguiente manera, en relación a cada sistema de clasificación diagnóstica respectivamente:

Tabla 8

Grupos de trastornos del DSM-IV-TR.

Trastornos que habitualmente se diagnostican por primera vez en la infancia o la adolescencia Delirium, demencia, trastornos amnésicos y otros trastornos cognoscitivos Trastornos mentales debido a una enfermedad médica Trastornos relacionados con sustancias Esquizofrenia y otros trastornos psicóticos Trastornos del estado de ánimo Trastornos de ansiedad Trastornos somatomorfos Trastornos facticios Trastornos disociativos Trastornos sexuales y de la identidad sexual Trastornos de la conducta alimentaria Trastornos del sueño Trastornos del control de los impulsos no clasificados en otro lugar Trastornos adaptativos Trastornos de la personalidad Otros trastornos que pueden ser foco de atención clínica
De American Psychiatric Association. <i>DSM-IV-TR. Manual diagnóstico y estadístico de los trastornos mentales. Texto revisado.</i> Barcelona: Masson, S.A., 2002, con autorización.

Fuente: Kaplan & Sadock. (2009). Sinopsis de Psiquiatría.

Tabla 9

Categorías de la CIE-10.

Trastornos mentales orgánicos, incluidos los sintomáticos. Trastornos mentales y del comportamiento debidos al consumo de sustancias psicoactivas. Esquizofrenia, trastorno esquizotípico y trastornos de ideas delirantes. Trastornos del humor (afectivos). Trastornos neuróticos, secundarios a situaciones estresantes y somatomorfos. Trastornos del comportamiento asociados a disfunciones fisiológicas y a factores somáticos. Trastornos de la personalidad y del comportamiento adulto. Retraso mental. Trastornos del desarrollo psicológico. Trastornos del comportamiento y de las emociones de comienzo habitual en la infancia y adolescencia.
De Organización Mundial de la Salud. <i>CIE-10. Clasificación Internacional de Enfermedades. Capítulo V.</i> 2000.

Fuente: OMS. (2000). CIE-10.

En su quinta versión, el Manual Diagnóstico y Estadístico de Trastornos Mentales (DSM-5) considera dentro del grupo de trastornos del estado de ánimo a todas aquellas morbilidades que conlleven un patrón de enfermedad derivado de un estado de ánimo alterado, en este grupo

incluye: “Trastorno depresivo mayor, trastorno depresivo persistente (distimia), trastorno de desregulación disruptiva del estado de ánimo, trastorno disfórico premenstrual, trastorno depresivo debido a otra afección médica, trastorno depresivo inducido por sustancias, otro trastorno depresivo no especificado, trastorno bipolar I y II, trastorno ciclotímico, bipolar inducido por sustancias, bipolar debido a otra afección médica y trastorno bipolar no especificado, como parte del reto diagnóstico incluyen trastornos que manifiestan síntomas depresivos y maníacos como al trastorno esquizoafectivo, trastorno de adaptación, algunos trastornos de la personalidad, proceso de duelo no complicado y otros trastornos” (Morrison, 2014, p. 110); cabe resaltar que a pesar de su nueva inclusión a la práctica clínica, la CIE-10 sigue siendo el sistema de clasificación más utilizado por el sistema de salud, por lo que su onceava versión se ha vuelto de las más esperadas.

Por lo que se ha señalado respecto a su distribución se hace posible observar el amplio rango discriminativo para el diagnóstico de estos sistemas de clasificación, lo cual ratifica su elección en algunos estudios de validación de pruebas diagnósticas como gold standard respecto a la capacidad discriminativa que otras pruebas pretendan delimitar, siendo este el caso demostrado por dos estudios de validación diagnóstica, el primero realizado en el año 2002 por Martínez, J. y su equipo de investigación de la Unidad clínica de Occidente en España, en la que adaptó y validó una versión abreviada del cuestionario de Yesavage para el despistaje de depresión en mayores de 65 años (GDS), en el que *“como parte de su procedimiento, un grupo independiente de investigadores realizó de forma ciega una entrevista semiestructurada para determinar síntomas de depresión según los criterios establecidos en la CIE-10, tomándola como la prueba de referencia para el diagnóstico”* (Martínez, Onís, Dueñas et al. 2002, p. 624); el segundo realizado el año 2006 en nuestro país, por De la Torre, J. y su equipo médico de investigación, los cuales realizaron una validación de la versión reducida de la escala de

depresión geriátrica (GDS) en el consultorio externo de geriatría del Hospital Nacional Cayetano Heredia, con un muestra de 400 pacientes mayores de 60 años que no recibían tratamiento antidepresivo ni tenían evidencia de deterioro cognitivo, a los *“que se les aplicó como gold standard criterios del DSM-IV, contrastada con la escala de depresión geriátrica abreviada”* (De la Torre et al. 2006), de ambos estudios se concluyó que la GDS abreviada ofrece *“una fiabilidad y validez diagnóstica similar a la escala completa, aunque sugieren que se tome como punto de corte 5 o más para incrementar su sensibilidad, sin embargo, que en términos prácticos sería una herramienta útil para el screening de depresión, además de reducir el tiempo de evaluación por paciente lo cual optimizaría la atención del mismo”* (Martínez et al. 2002; De la Torre et al. 2006).

Algunas otras formas de aproximarse al diagnóstico de depresión es a través de los resultados obtenidos por la aplicación de herramientas clínicas, como inventarios, escalas, etc. algunos recomendados como pruebas a consideración como parte de la exploración mental o examen psicopatológico, entre algunas escalas de puntuación objetiva para la depresión los expertos sugieren a la Escala de autoevaluación de depresión de Zung, la cual proporciona un índice global de la intensidad sintomatológica, así mismo se sugiere la Escala de depresión de Raskin la cual mide también la gravedad del cuadro depresivo según los informes del paciente contrastados con los del especialista y finalmente es importante considerar la Escala de depresión de Hamilton, la cual evalúa diversas dimensiones que constituyen la manifestación sindrómica de la depresión (Kaplan y Sadock, 2009); de entre estas escalas sólo dos registran estudios que valoran su capacidad discriminativa para el diagnóstico de la depresión, a saber, la Escala de depresión Zung y Hamilton, habiéndose ya descrito algunos estudios de la primera escala a diferencia de la de Hamilton, de la que Montenegro, M. en el 2014 en su estudio de utilidad del PH-9 en el diagnóstico de trastorno depresivo considera a la Escala de depresión

de Hamilton como gold standard en Argentina para valorar trastornos depresivos (Montenegro y Kilstein, 2014), recomendando un entrenamiento en su utilización para evitar sesgos que afecten su discriminación diagnóstica, por otro lado, Heinze, G. respecto a la determinación numérica para la severidad de la depresión sugiere también la utilización de la escala de Hamilton (Heinze y Camacho, 2010), pese a ello sugiere más estudios que incrementen su evidencia ya que a la fecha se cuenta con evidencia baja acerca de su efectividad.

De entre las muchas escalas sugeridas por algunos estudios de validación, la Escala de Zung para evaluar depresión resulta una de las que demuestra mayor número de estudios que sustentan su utilización como prueba de referencia (incluye los estudios del marco referencial); en nuestro medio (contexto de estudio: Ciudad de Huancayo; Junín-Perú) se dispone a esta escala como una de las más usadas al momento de complementar hallazgos de la exploración clínica, ha sido utilizada como prueba de primera línea para la determinación cuantitativa de la sintomatología depresiva tanto para la práctica clínica como de la salud.

La escala de autoevaluación de la depresión (EAMD) se encuentra destinada a la estimación numérica o cuantitativa de la depresión como condición clínica (síntoma o signo presente en una amplia gama de trastornos clínicos), tanto la EAA como la EAMD fueron publicadas por primera vez en 1965, la EAMD también compuesta por 20 reactivos, en la que un índice bajo indicaría la inexistencia de depresión significativamente clínica y un puntaje al extremo opuesto indicaría la presencia de depresión de significación clínica, para la cual también se establecieron rasgos de corte para su severidad (Badillo, 2017), las cuales se presentan como sigue (las categorías diagnósticas fueron homologadas por razones prácticas):

- Depresión normal (<40).
- Depresión mínima a moderada (41-47).
- Depresión marcada a severa (48-55).
- Depresión en grado máximo (56>).

La EAMD evalúa la depresión bajo 4 dimensiones básicas, a saber, reactivos distribuidos en dimensiones como estado afectivo persistente, fisiológicos, psicomotor y psíquico.

Respecto a las pruebas de estudio el Inventario de Depresión de Beck (BDI) tiene como objetivo tras su segunda adaptación a criterios DSM-IV perseguir el diagnóstico de depresión, además de ajustarse al modelo de depresión de Beck, consta de 21 reactivos que evalúan un amplio espectro de síntomas depresivos, compuesto por reactivos en 4 niveles de afirmación, evaluando de entre sus múltiples sub-dimensiones el sentimiento de culpa, la pérdida de apetito, preocupación somática, etc. Considerándose una de las pruebas con mayores utilidades en estudios científicos, su aplicación a partir de los 13 años, precisa para personas con nivel educativo medio-bajo con un tiempo corto de administración, como parte de su construcción y adaptación en su nueva versión el BDI-II cuenta con valores de discriminación diagnóstica, la cuales le obtuvieron a través del estudio de la prueba en muestras por casos y controles, diagnosticando antes a los grupos por un equipo de trabajo sobre la base de los criterios del DSM-IV, para que luego sean clasificados y luego del análisis del *test post hoc de Scheffé* indicó puntuaciones considerables [$p < 0,001$], lo cual indicaría que el BDI-II es una herramienta válida para discriminar a los pacientes con trastornos del estado de ánimo de aquellos que presentan otros síndromes, la utilización de curvas ROC se realizó para establecer las puntuaciones óptimas de corte para los rangos que determinan su severidad (Beck, Steer y Brown, 2009, p. 82), de estas se obtuvo la siguiente clasificación diagnóstica:

- Depresión mínima (<13).
- Depresión leve (14-19).
- Depresión moderada (20-28).
- Depresión severa (29>).

Además del análisis de su capacidad diagnóstica también los autores han realizado validez de contenido, estabilidad por test re-test, y sobre todo el análisis factorial de sus dimensiones, las cuales también ofrecen valores significativos de aceptación.

Del inventario clínico podemos advertir que se dispone como una herramienta que además de actualizada en su marco teórico, también se hallan adaptadas a criterios diagnósticos DSM, por lo que sus valores clínicos aunados al material en evidencia de precisión diagnóstica nos otorga mayor oportunidad como para considerarlas bajo un estudio de rigor científico como el de la validación diagnóstica en las nuevas pruebas de referencia en lo que la práctica clínica y de la salud corresponde en la región, respecto a la prueba ya existente como gold standard.

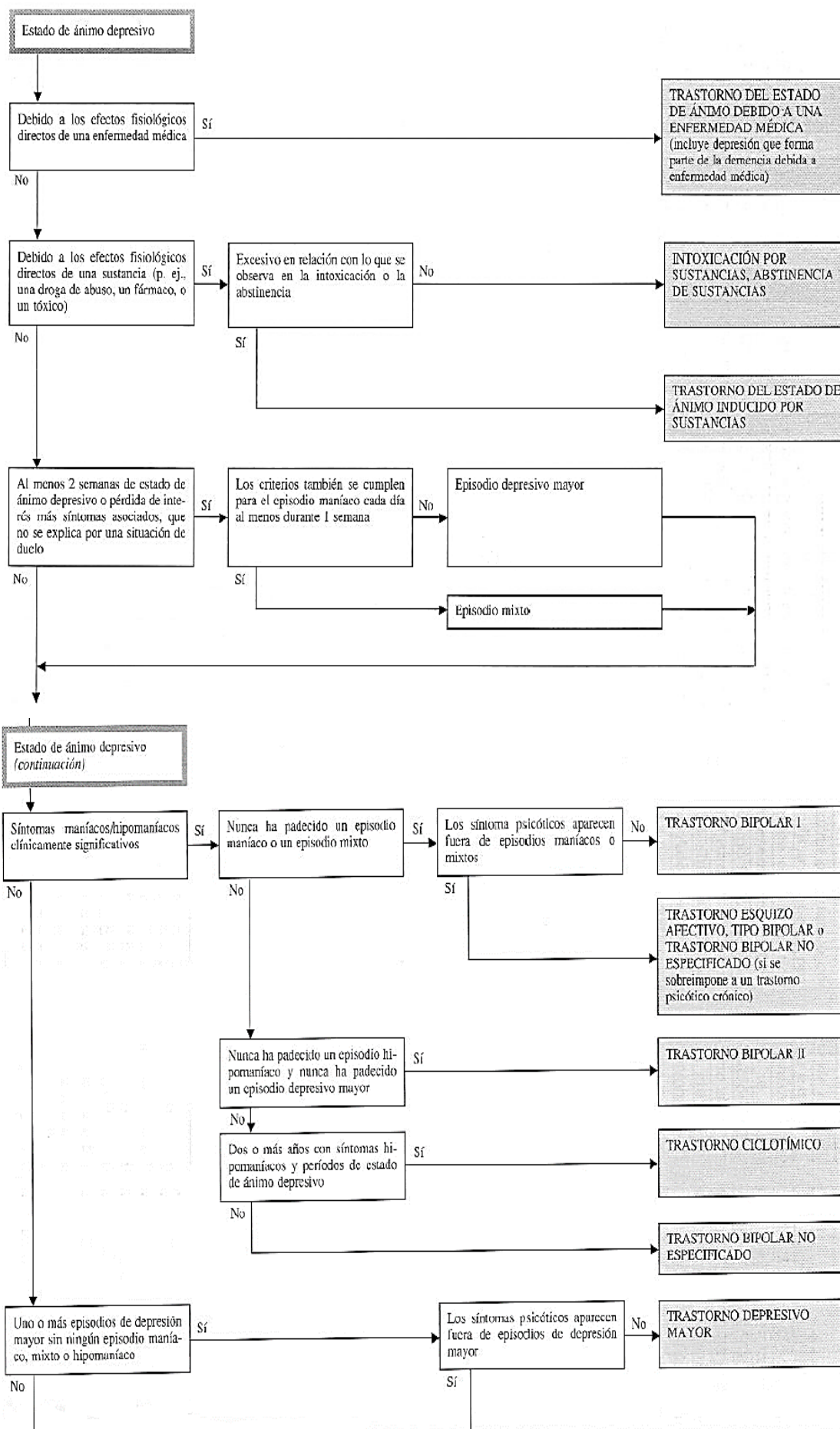
Diagnóstico Diferencial:

Los procedimientos concatenados a la formulación diagnóstica de la condición de salud-enfermedad de algún paciente traen subyacentemente la determinación de diagnósticos diferenciales respecto a toda la manifestación sintomatológica que exhibe y su contigua delimitación en un trastorno específico del vasto universo de enfermedades exploradas en nuestra especialidad, por lo que su elaboración constituye el pilar del trabajo clínico habitual, por lo que la recogida de datos resumidos los registros sistemáticos que el clínico pueda disponer en su práctica, facilitan el primer paso para discernir entre el diagnóstico que se ajuste

mejor a la realidad que expone cada paciente (First, Frances y Pincus, 2006), dicha capacidad discriminativa también se encuentra apoyada por las pruebas diagnósticas, ya que recopilan información traducida en categorizaciones numéricas del estado del paciente las cuales en su amplia diversidad integren consistencia científica al proceso de exploración clínica.

La experiencia de malestar expresada por el paciente, denominada síntoma y la consecutiva transcripción clínica a signos, nos permiten filtrar las posibilidades esencialmente bajo 6 caminos diagnósticos, para descartar, en primer lugar, “*simulación/trastorno facticio*, en segundo lugar, descartar una *sustancia como etiología*, luego descartar una *etiología médica*, en cuarto lugar, determinar los *trastornos primarios específicos de los comórbidos*, posteriormente *diferenciar un trastorno adaptativo de un trastorno no específico* y finalmente *delimitar la frontera con la ausencia de trastorno mental*” (First, Frances & Pincus, 2006).

En este sentido la estructuración de árboles de diagnóstico diferencial sistematiza el conglomerado de criterios diferenciales en prácticas rutas diagnósticas, representan por lo tanto un método de enrutamiento clínica, a modo de algoritmos o a modo de esquemas, orientan al clínico desde la presentación de un síntoma en específico y proporciona puntos de decisión para determinar qué diagnóstico sería el adecuado (First, Frances y Pincus, 2006), a continuación se presenta el árbol de diagnóstico diferencial respecto a la depresión.



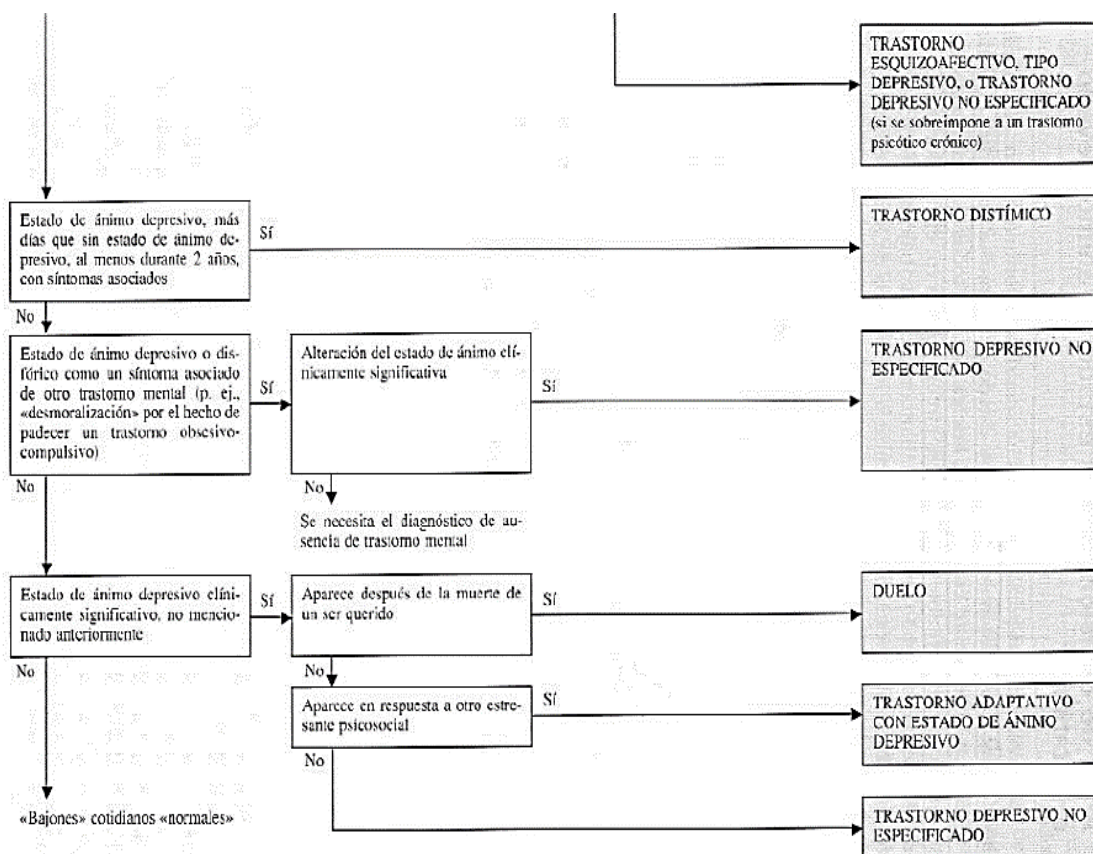


Figura 4 Árbol de decisión para el estado de ánimo depresivo.

Fuente: First, Frances & Pincus. (2006). Manual de diagnóstico diferencial (p. 67).

MEDICINA BASADA EN LA EVIDENCIA (MBE):

Se hace imposible describir la Medicina Basada en la Evidencia, su recuento histórico, principios etc. sin antes hablar de la práctica médica, la cual se ve fundamentada en parte gracias a la MBE, por lo que su práctica se fundamenta en todo el avance científico que han desarrollado sus profesionales desde hace más de 60 años, propiciando la expansión de su información científica, modificando de manera transcendental las formas de abordar la enfermedad misma (con lo que todo este concepto incluye), manteniendo de esta forma su carácter imparale para generar nuevos y mejores conocimientos por cada generación de profesionales que la practican (Harrison, Longo, Fauci et al. 2012, p. 2).

Por lo que la formación educativa demanda la prioridad de incidir favorablemente sobre la adquisición de habilidades y la formación de actitudes (de carácter científico), con la finalidad de ir consolidando la competencia clínica y hacer de cada estudiante un profesional en la capacidad autónoma de generar conocimientos de rigor científico productivos para su propia ciencia y sociedad, actividad claramente presidida por el pensamiento central: el bienestar de nuestros pacientes. “Pensamiento contextualizado en su práctica gracias a la utilización de métodos científicos para resolver problemas, acompañada de comprensión humana y sabiduría en beneficio propio y de la sociedad” (Farreras & Rozman, 2012, p. 3).

Es pues de entre muchas actividades que constituyen la práctica médica, las habilidades clínicas las que se relacionan con la presente investigación, de esta, en especial para la elaboración del diagnóstico, enmarcado en lo que se denomina como método clínico (relacionado al método científico en un sentido pragmático), en el que su proceso sigue una serie no siempre concatenada de pasos (pero en la posibilidad del médico o profesional de la salud, lógica) a partir de la obtención de datos del paciente, dicha recolección de datos (exploración clínica) ha venido mostrando mejoras respecto a sus métodos y principios fundamentados en investigación científica, pasando de la exploración física u ordenación sindrómica a la utilización de herramientas complementarias (pruebas diagnósticas) bastante sofisticadas en función a su tecnología o calidad de datos como complemento clínico al momento de tomar decisiones (Harrison et al. 2012; Farreras et al. 2012).

La suma de estas u otras estrategias clínicas conforma lo que hoy en día conocemos como método clínico, y al respecto de las pruebas complementarias es preciso entenderlas como tal, es decir, “pruebas complementarias” (énfasis añadido) al procedimiento clínico, el cual es exclusivamente definido por el juicio o valoración del médico o profesional de la salud, sin

embargo, lo que ameritaría su elección serían exploraciones completarias que sean estrictamente necesarias a la naturaleza del estudio de caso en cuestión (p. ej. atipicidad, alto grado de incertidumbre, etc.), esto requiere a su vez de tener en cuenta las cualidades diagnósticas (parámetros de precisión y calidad diagnóstica) de diversas pruebas, específicamente guiándonos del grado de sensibilidad y especificidad que ofrece cada una (Farreras & Rozman, 2012, p.5).

Por lo que es puntual que el profesional en todo momento del proceso diagnóstico, tenga claridad de la jerarquía (*jerarquía en las pruebas complementarias*) (Farreras y Rozman, 2012, p.5) de las pruebas elegidas para el apoyo a su exploración, en el sentido de la evidencia de la cual dispone dicho instrumento elegido y si sus parámetros son lo suficientemente significativos como para utilizarla respecto a otra prueba de referencia (fundamento científico), lo cual implicaría además de una mejora en la exploración para el paciente, también implicaría menores costes para el sistema de salud por la utilización conglomerada de múltiples procedimientos diagnósticos complementarios. Por otro lado, Farreras (2012) aclara que su interpretación también debería ser pertinente, en función a los resultados de la misma prueba y a la significación de estos de acuerdo a los parámetros de discriminación diagnóstica, ya que una prueba complementaria no sólo es una serie de reactivos y respuestas, sino todo un procedimiento cuantitativo de valoración objetiva de las manifestaciones clínicas de alguna enfermedad, condición de salud, etc. Resumidamente, el profesional de la salud será responsable inmediato de que las exploraciones complementarias constituyan un gradiente coherente con el proceso presidido por el ejercicio clínico (p. 6), es así que al hablar de jerarquías de pruebas complementarias o evidencia como respaldo científico, podemos hablar de Medicina Basada en la Evidencia.

La *Medicina Basada en la Evidencia* fue desarrollada por un equipo de médicos internistas y epidemiólogos, sus antecedentes históricos recientes ratifican que inició en 1992, presidido principalmente por Gordon Guyatt, aunque para el siglo XIX Bichat y Magendi describieron sus primeros fundamentos filosóficos; la cual supone la integración de la experiencia clínica acumulada asociada a la mejor evidencia disponible derivada de la investigación científica sistemática, siendo enfocado su ejercicio no sólo a la práctica médica sino a todo profesional que tiene posibilidad de intervención o de tomar decisiones en las ciencias de la salud, por lo que supone considerarla también como una serie de actitudes hacia la práctica y aprendizaje de las ciencias ligadas a la salud (Ybarra et al. 2015; Arceo et al. 2010).

En un estudio realizado por Sánchez, M. (2001) describe el proceso que conlleva realizar alguna práctica en salud en el marco de la MBE, la cual la estructura en 5 pasos consecutivos (p. 99), las cuales se muestran, como sigue:

- *Primer paso:* elaboración de una pregunta clínica relevante y bien estructurada (proveniente de un problema clínico en contexto real).
- *Segundo paso:* esfuerzos orientados a la búsqueda sistemática de la mejor evidencia científica actual, la cual se delimita por el mismo problema clínico y no por rutinas, hábitos o protocolos individuales.
- *Tercer paso:* screening, como análisis crítico de la información encontrada.
- *Cuarto paso:* utilización de la información evaluada (filtrada) en la solución del problema clínico.
- *Quinto paso:* se reevalúa todo el proceso para aportar utilidad al procedimiento clínico.

Con esto se propone dar la mejor utilidad a la mejor evidencia dispuesta en el campo de la medicina, sin desconocer la importancia de la experiencia, aludiendo con mejor evidencia a los estudios sujetos a diseños de significativo rigor científico, de la que se ha elaborado con el avance de este campo del conocimiento una jerarquía de delimita su valor científico, dicha escala va del primer nivel de calidad (Revisión sistemática y meta-análisis: I) el cual ofrece un mínimo sesgo de resultados hasta el último nivel, el cual lo constituye sin duda la Experiencia clínica [IV] el cual por su naturaleza de adquisición ofrece el grado máximo de sesgo (Ybarra, Orozco y Valencia, 2015), ya que es producto de la observación o especulación no controlada científicamente.

Estos criterios aportan sin duda alguna, ventajas al proceder clínico, de entre ellas a la disposición de formular conocimiento actualizado que mejore su juicio crítico, incremento de la confianza de las acciones o procedimientos médicos, debido a la búsqueda y análisis de la información científica, sumando base científica a los criterios de conocimiento, lo cual mejora en términos generales la práctica médica llevándola gradualmente a su super-especialización. En otro sentido de las cosas, las ventajas también se proyectan a los pacientes, disminuyendo por ejemplo la variabilidad de conocimiento a los que se pueden ver expuestos (como parte de la educación sanitaria preventiva o promocional), mejorando la relación médico-paciente, tornándola activa en ambos sentidos, además de ser una manera de democratizar el conocimiento impartido lo cual orienta a una atención integral, actualizada y documentada (Arceo, Ornelas & Domínguez, 2010).

PSICOLOGÍA CLÍNICA BASADA EN LA EVIDENCIA (PCBE):

Desde el estudio de eficacia terapéutica realizada por Eysenck, hasta la inclusión de la nueva generación de terapias contextuales para diversos trastornos mentales, la Psicología Basada en la Evidencia lleva siendo ya practicada varias décadas, por lo que su surgimiento más elemental se remonta a la creación de la primera clínica psicológica gracias a los esfuerzos de Lightner Witmer en 1896 en la que se basaba la práctica clínica de la psicología en evidencia científica producto de continuas investigaciones, ya que el descubrimiento de nuevos trastornos y su crecimiento epidemiológico fortaleció la necesidad de desarrollar modelos teóricos explicativos y aplicativos (terapéuticos) fundamentados en la investigación, gracias a esto fue también que las prácticas psicoterapéuticas basadas en intuición o sistemas meramente especulativos se encuentran siendo descartadas actualmente por la comunidad científica, como se mencionó anteriormente Eysenck en 1952 realizó un estudio en el que comparó la eficacia de los tratamientos psicológicos en muestras de casos y controles, a los que se les realizó la intervención durante 6 meses y de la que se halló que no hubo diferencias significativas entre ambos grupos, lo cual generó debate en la comunidad profesional de la época y que en respuesta a dicho descubrimiento fundó un laboratorio experimental del comportamiento animal el cual aún se encuentra en funcionamiento y sigue realizando investigaciones a favor de los mecanismos que generan realmente el cambio en procesos patológicos. Para posteriores años se comenzó a encontrar evidencia de mayor eficacia terapéutica, por lo que comenzaron a encontrar que el recibir tratamiento psicológico generaba efectos modificadores de la salud afectada o mostrándose mejor que el placebo y una vez demostrada la eficacia de la terapia psicológica sobre el placebo el debate cambió de orientación, enfocándose a la efectividad entre diversas terapias psicológicas, de la que gracias diversos meta-análisis la terapia de conducta demostró la superioridad terapéutica específicamente en algunos trastornos, datos que fueron

verificándose y generalizándose estudios para otros trastornos menos específicos (Ybarra et al. 2015).

Producto de estas iniciativas en 1993 la APA creó la división 12 la cual se aportó entre muchas cosas los criterios que delimitan con claridad los tratamientos psicológicos con validez empírica y por lo tanto las que representan las primeras líneas de elección terapéutica, categorizándolos en 1998 según su calidad y rigor científico en básicamente 3 grupos de tratamientos, siendo este trabajo el primero que también expuso por primera vez la distinción conceptual entre eficacia, efectividad y eficiencia de algún tratamiento psicológico.

Tabla 10

Criterios de validez para tratamientos psicológicos.

Tratamiento bien establecido (condiciones I o II)
<p>I. Al menos dos estudios con diseño de grupo llevados a cabo por investigadores diferentes deben haber demostrado la eficacia del tratamiento en una o más de las siguientes maneras:</p> <p>a) El tratamiento es superior al tratamiento farmacológico o al placebo, o a otro tratamiento.</p> <p>b) El tratamiento es equivalente a otro establecido en estudios con adecuado poder estadístico.</p> <p>II. Demostrar eficacia mediante una serie amplia de estudios de caso único. Estos estudios deben tener:</p> <p>a) Un buen diseño experimental.</p> <p>b) Haber comparado la intervención dentro del modelo de diseño único con el placebo, con la intervención farmacológica o con otros.</p>
Tratamiento probablemente eficaz (condiciones III o IV o V o VI)
<p>III. Dos estudios han mostrado que el tratamiento es más eficaz que un grupo control en lista de espera.</p> <p>IV. En dos estudios se cumplen los criterios de tratamiento bien establecido pero son llevados a cabo por el mismo investigador.</p> <p>V. Al menos dos buenos estudios demuestran la efectividad del tratamiento pero muestran una gran heterogeneidad en las muestras de clientes.</p> <p>VI. Usando metodología de caso único se reúnen los criterios para tratamiento bien establecido, pero la serie de estudios es pequeña (menos de tres).</p>
Tratamiento en fase experimental
<p>VII. Tratamientos que no se pueden enmarcar en las categorías anteriores, aunque existan indicios experimentales prometedores de que se pueden alcanzar los niveles de eficacia establecidos.</p>

Fuente: Pérez et al. (2003). Guía de tratamientos psicológicos eficaces I Adultos (p. 27).

En este marco, las investigaciones clínicas que confieran una gran validez interna permitirán determinar el alcance o impacto terapéutico de algún modelo de intervención (demostrando su grado de eficacia), así como también permitirán saber si los cambios generados están sujetos a los componentes del tratamiento y no a otros elementos producto de sesgo (demostrando su grado de efectividad) y finalmente su generalización de resultado a contextos reales, esperando

que las respuestas experimentales demuestren consistencia y estabilidad en situaciones reales lo cual demostraría su grado de eficiencia.

Ulteriormente se ha visto que esta actitud de fundamentación científica se ha ido extrapolando a otros campos de la psicología más que sólo enfocándose en la calidad terapéutica en el área clínica, sino también en el área de la salud e incluso para el desarrollo y gestión de estrategias en salud pública, pero al parecer no se ha integrado rotundamente para las habilidades clínicas específicas del psicólogo, como la evaluación, diagnóstico, prevención, etc. (prácticas en salud basadas en la evidencia), campo pendiente en la investigación.

TERAPIAS PSICOLÓGICAS BASADAS EN LA EVIDENCIA:

En la línea de la MBE o la PCBE, las prácticas psicoterapéuticas realizadas por mera especulación o intuición de funcionamiento son el principal reto de erradicación, por lo que en las últimas décadas las terapias psicológicas se han habituado a la validación empírica de sus principios, gracias a la integración de las investigaciones científicas ligadas a la práctica clínica fundamentada en la mejor evidencia en materia de estudios de alto rigor metodológico, por lo que hemos pasado de determinar la eficacia de alguna terapia para algún trastorno específico a estimar la proporción de eficacia, eficiencia y efectividad entre terapias para distintas condiciones y saber así cuales son las más apropiadas como primera elección terapéutica (Echeburúa, Salaberría, Corral et al. 2010).

Sólo para tener en cuenta la relevancia de incluir intervenciones apoyadas en evidencia, “el registro realizado por Garfield en 1960 demostraba que para aquel entonces se contaba con al menos 60 formas de psicoterapia y que para las décadas de los 70s y 80s al menos existían 130 tipos más, tendencia verificada por Kazdin el cual encuentra la convivencia de entre al menos

400 terapias en la práctica clínica, tendencia que ha venido creciendo con el paso del tiempo y la falta de formalización científica en la Psicología, por lo que se espera que esta tendencia al menos se duplique el próximo siglo, tendencia lamentable ya que nuestros avances terapéuticos no se están desarrollando alternadamente al avance científico” (Vera, 2004, p. 6), por lo que los antecedentes de efectividad de terapias psicológicas son carentes incluso ahora cuando se ha producido una forma de revolución para la producción de conocimientos en nuestra especialidad. En respuesta a dicha problemática muchos grupos académicos respetables han comenzado a fortalecer esta otra nueva tendencia de analizar las intervenciones terapéuticas y aportarles soporte empírico gracias a la investigación de alto rigor científico y control de sesgo, intervenciones que se muestren más eficaces que el placebo, la no intervención o el tratamiento psicofarmacológico.

Tal intención ha ido ampliando sus alcances también a la formación académica profesional y sobre todo ha incidido en las maneras en las que se dispone la formación a estudiantes o la manera en la que se les suministra contenidos de información, los cuales deberían ser en primera instancia la mejor evidencia científica con la que disponemos en la actualidad, aspecto que se ha comenzado a reflejar en pequeña proporción en la práctica clínica habitual de algunos profesionales formados en calidad de la actitud científica (Echeburúa y Corral, 2005). Es por ello que la evaluación de la eficacia de prácticas o intervenciones terapéuticas son una necesidad para nuestra ciencia por los siguientes motivos determinados por Barlow y Hoffman en 1997:

- *En primer lugar:* Porque la delimitación de procedimientos terapéuticos eficaces se hace implícito como actividad propia del progreso científico de la Psicología.

- *En segundo lugar:* Porque la disciplina se enfrenta a retos clínicos de mayor complejidad en función de su propio desarrollo.
- *En tercer lugar:* Porque de este modo las terapias ofrecidas en instituciones públicas, al servicio de la mayor cantidad de la población clínica de cualquier país tendrán la mejor calidad respecto a otros procedimientos no basadas en la evidencia, fortaleciendo nuestro nexo con las poblaciones y encontrando el reconocimiento de valor científico por los mismos.

En este sentido, actualmente como parte de la evaluación de eficacia de procedimientos terapéuticos se ha establecido gracias a la Agency for Health Care Policy and Research (AHRP) niveles de evidencia que usualmente muestran algunos estudios de evidencia, niveles que serán tenidos en cuenta al momento de la elección de una formación terapéutica o en su defecto de la elección de alguna estrategia terapéutica en la práctica clínica diaria.

Tabla 11
Niveles de evidencia.



NIVEL DE EVIDENCIA	TIPO DE ESTUDIO
Mínima	<ul style="list-style-type: none"> • Descripción de casos únicos. • Estudios descriptivos o resoluciones de comités de expertos. • Estudios de caso controlados. • Ensayos controlados bien diseñados, pero no aleatorizados, o estudios naturalistas. • Ensayos controlados y aleatorizados con muestras pequeñas. • Ensayos controlados y aleatorizados con muestras grandes. • Meta-análisis de ensayos controlados y aleatorizados con muestras suficientemente grandes.
Máxima	

Fuente: Echeburúa et al. (2010). Terapias psicológicas basadas en la evidencia (p. 248).

Por lo mismo se recomienda apearse a criterios para el éxito en la terapia, relacionados sobre todo tanto a la validez interna, por la cual en la medida de su magnitud el clínico sabrá que el cambio terapéutico se debió no a otros elementos más que a los de la terapia utilizada y

consecuentemente al utilizarla en poblaciones diversas valorar su efectividad por el grado de generalización de resultados en otros contextos similares, pero no iguales al aplico en principio.

Tabla 12
Criterios de éxito en la terapia.

	CONCEPTOS FUNDAMENTALES	DEFINICIÓN	
Validez interna 	Eficacia	Resultados favorables en un contexto de investigación.	 Validez externa
	Efectividad	Resultados favorables en la práctica clínica habitual.	
	Eficiencia	Buenos resultados al menor coste posible.	

Fuente: Echeburúa et al. (2010). Terapias psicológicas basadas en la evidencia (p. 251).

Terapéutica de Variable Clínica de Estudio Diagnóstico:

A la fecha algunos estudios de meta-análisis consideran a las siguientes terapias como las de primera línea en lo que evidencia respecta para tratar trastornos de ansiedad o del estado de ánimo, pero cabe resaltar la superioridad de las terapias conductuales, cognitivas y contextuales por la capacidad que han demostrado para trastornos específicos y frente a otras terapias existentes; en nuestro país de entre muchos estudios de eficacia terapéutica con niveles de significancia del al menos $p < 0.05$ evidencian que las intervenciones cognitivas conductuales aplicadas a adolescentes con sintomatología depresiva, en el cual se evaluó los síntomas mediante el BDI como herramienta pre y post prueba a la intervención (*su utilización implica aportar parámetros de validez diagnóstica para otorgar mayor consistencia a los resultados, siendo el BDI o BAI pruebas muy utilizadas como herramientas de monitorización en este tipo de estudios*), demostrando su efectividad para la reducción sintomatológica, en otro estudio se comprobó que cierto número de sesiones cognitivas conductuales redujeron notablemente la sintomatología, demostrando su análisis estadístico un $p < 0.004$ (Vera, 2004, p. 7) como nivel

de significancia estadística y entre otros estudios que demuestran niveles de significancia importantes como para concluir que la terapia cognitiva conductual demuestra eficacia terapéutica.

De tal manera se presentan las terapias basadas en la mejor evidencia clasificadas según el grado de rigor científico por las cuales fueron analizadas en el 2001; a continuación, se presenta una lista breve en lo que depresión respecta:

Tabla 13

Terapias psicológicas apoyadas por evidencia (depresión).

3. Depresión	
Depresión mayor	-Terapia conductual: I
	-Terapia dinámica breve: II
	-TCC: I
	-Terapia interpersonal: I
	-Terapia de autocontrol: II
	-Resolución de problemas sociales: II

Fuente: Ybarra, Orozco & Valencia. (2015). Intervenciones con apoyo empírico
Herramientas fundamentales para el psicólogo clínico y de la salud (p. 14).

Del cuadro anterior podemos evidenciar que cada estrategia terapéutica que presenta alguna forma de estudio de respaldo como evidencia tiene al lado derecho I, II o en pocos casos III, estos hacen referencia a la calidad de evidencia que demuestran, siendo I: Tratamientos bien establecidos, II: Tratamientos probablemente eficaces y III: Tratamientos en fase experimental.

Es oportuno agregar que la importancia de aportar estudios que respalden la evidencia de pruebas clínicas utilizadas ya sea como estrategias complementarias al diagnóstico o como herramientas de verificación del impacto de un procedimiento terapéutico (herramienta de monitorización o comprobación pre y post tratamiento) como parte del avance científico en nuestra ciencia. Actualmente, por ejemplo, muchos de los manuales de terapia cognitiva

conductual agregan como parte de sus procedimientos de formulación clínica la inclusión de herramientas como el Inventario de Depresión de Beck (al igual que como se utiliza los SUDS o USAs, como escala de unidades subjetivas de perturbación o ansiedad) para valorar la sintomatología inicial, facilitando la elaboración de líneas base de comprobación (Caballo, 2002, p. 223), sobre todo para trastornos del estado de ánimo, por lo que tal propósito se vería complementado consistentemente si las pruebas elegidas también cuentan con algún nivel de evidencia, el cual pueda además acreditar que dichos sesgos respecto a parámetros del impacto de la terapia o de la misma prueba fueron controlados con rigor científico.

Caso análogo para algunas terapias de tercera generación y, caso especial para la Activación Conductual como terapia psicológica respaldada por evidencia que sugiere su grado de efectividad muy significativa respecto a la Terapia Cognitiva para la depresión o incluso la misma Terapia Cognitiva Conductual; “volviéndola parte de la elección de primera línea terapéutica, la cual como parte de su proceso terapéutico precisa como primera actividad la formulación de una línea base de contrastación y con ello la elección de alguna prueba clínica que sirva como herramienta de monitorización”, y en estudios de caso clínico aplicando este modelo terapéutico se sugieren utilizar BDI, Escala de Zung o Hamilton como herramientas de seguimiento y contrastación de evolución, priorizando por supuesto la utilización del BDI por sobre las otras pruebas por su relativa contemporaneidad y modelo teórico más relacionado al modelo terapéutico en cuestión (Lejuez, Hopko y Hopko, 2001, p. 264); sin embargo, lo que no compilan se relaciona a valores de precisión diagnóstica para las pruebas sugeridas, lo cual actuaría en el peor de los casos como variable modificadora del efecto mismo que implica la evaluación y recolección de datos o respecto al impacto real que está generando dicho procedimiento terapéutico, aspecto a tener en cuenta para el diseño y ejecución de estudios de eficacia terapéutica.

Pronóstico:

Cuando hablamos de pronóstico nos referimos a la capacidad predictiva o estimación de predicción que demuestra el profesional de la salud tras la evaluación y diagnóstico clínico, por lo que el pronóstico se constituye en sí mismo como una hipótesis de la posible evolución clínica a posteriori que cursará el paciente tras la administración de intervención clínica (Llanio y Perdomo, 2003), la cual plantea siempre 2 cuestiones esenciales respecto al paciente, la primera concerniente a la amplitud de la enfermedad (alcance e impacto de la enfermedad), es decir, qué tanto compromete la vida del enfermo y la segunda, los efectos adversos, rezagos o consecuencias que trae consigo; sumado a ello, el pronóstico representa también la manera en que los factores sociodemográficos congenian con la enfermedad, por lo que aquí importa variables como edad, sexo, condición socioeconómica, contexto de interacción, soporte familiar y otras muchas variables que configuran también la posibilidad de estimar un pronóstico favorable, desfavorable o algunos casos inciertos.

En el caso de la depresión su pronóstico es con frecuencia favorable si los pacientes son diagnosticados preventivamente y no cuando el paciente estuvo manteniendo la enfermedad durante muchos años, a pesar de que su incidencia para la cronicidad sea alta la depresión suele remitir en cuestión de meses con tratamientos conductuales, la prevención de sus recaídas son favorables en su mayoría, aunque la mayor complicación aún no controlada efectivamente radique en episodios graves con intentos de suicidio (Alarcón, Mazzotti & Nicolini, 2005).

Limitaciones Clínicas Para el Pronóstico:

La actividad predictiva es con mucha probabilidad una de las mejores habilidades clínicas producto inmediato de la actividad científica en la práctica clínica, su complejidad acompañada de aptitudes necesarias para su adecuada formulación predictiva la ha llevado a trazar sus

limitaciones respecto a la evaluación que realiza el clínico, su función y la calidad en su procedimiento, este último fundamentada en los conocimientos adquiridos como parte de la evaluación del problema (estudio de caso), iniciando por las ordenaciones sindrómicas (delimitación de signos y síntomas), hasta realizar estudios longitudinales de la historia del paciente para definir su diagnóstico, por lo que una buena formulación pronóstica refleja una buena parte del cuadro clínico, la evaluación diagnóstica y el tratamiento; es justamente en uno de estos tres últimos elementos mencionados que se relaciona la problemática de investigación con la limitación clínica para la formulación del pronóstico, a saber, la evaluación diagnóstica, y, es que en la evaluación psicológica se ha erradicado la atención que merece el pronóstico como elemento determinante para cualquier proceder clínico-terapéutico, realidad antagónica en el campo de la medicina, puesto que los principales inventarios, manuales, escalas, etc. dirigidos a la evaluación clínica de síntomas en nuestra especialidad minimiza e incluso ignoran su utilidad como parte del proceso pronóstico, por tanto, el clínico limitará su capacidad de abordaje al no estar apegado a conocimientos especializados de su profesión que le permitan suministrar mayor validez a su evaluación clínica, tratándose el pronóstico de un fenómeno naturalmente probabilístico, la utilización correcta de las pruebas clínicas sujetas a parámetros de precisión diagnóstica favorecería siempre que sea posible una formulación predictiva de calidad (Ibáñez y Echeburúa, 2015, p. 19), relacionada no sólo a las habilidades cognitivas o metacognitivas por parte del clínico, sino también validando su procedimiento por la inclusión de valores predictivos como parámetros de decisión diagnóstica.

INVESTIGACIÓN CLÍNICA PARA EL DIAGNÓSTICO:

Las revisiones sistemáticas de literatura académica necesarias para la presente han evidenciado notoriamente que los estudios observacionales y estudios de eficacia terapéutica son los realizados con mayor frecuencia dentro del amplio campo metodológico de la

Investigación Clínica, contradictoriamente los estudios de validez o precisión diagnóstica que son los realizados con menor frecuencia, a pesar de que en el caso de los estudios observacionales es la que demuestra mayor complejidad metodológica (mayor rigor científico) por lo que aporta mayor evidencia para el diagnóstico y en el caso de los estudios de eficacia terapéutica por la dependencia que estos últimos muestran por una elaboración correcta del diagnóstico previo a la exposición terapéutica con la que se estimará el impacto real de los componentes efectivos del tratamiento (Manterola, 2009; Moriana & Martínez, 2011).

Siendo el diagnóstico considerado como la actividad clínica central respecto al tratamiento y pronóstico, ya que del diagnóstico dependen los otros dos procesos en la atención del paciente, considerándola piedra angular de la medicina o de toda práctica definitiva en las ciencias de la salud, representando los esfuerzos técnicos (la elección y administración de pruebas clínicas complementarias) una parte importante del soporte científico que el clínico le puede dar a su proceso de determinación diagnóstica, traducido en su juicio clínico para firmar o negar la existencia de enfermedad (Llanio y Perdomo, 2003), condición patológica, sintomatología perpetuante o en términos generales alguna afectación de la salud integral del individuo. “Dicho juicio acompañado de imposibilidad de ser certera completamente, ya que el objetivo del especialista no es alcanzar la certeza absoluta del diagnóstico sino reducir suficientemente los niveles de incertidumbre como para tomar decisiones terapéuticas lo más confiables posibles” (Fagan, 1975, p. 258), aspecto que puede ser alcanzado tanto por las fuentes de información (datos epidemiológicos de cada paciente) recolectadas y las herramientas complementarias que contribuyen el soporte cuantitativo de nuestras observaciones posiblemente sesgadas.

El presente estudio de precisión diagnóstica se formula en contextos clínicos (el problema de investigación se halla en la práctica clínica cotidiana; al generar ideas respecto al *diagnóstico*, tratamiento, pronóstico, etc.), por lo que su formulación problemática y posterior manejo metodológico para dar resolución científica más oportuna al problema (Equivalente al acrónimo *PICO* utilizado en medicina basada en evidencias, Pacientes: Adultos; Intervención: Diagnóstico; Comparación: Estudio de precisión diagnóstica; Outcomes o Resultados: Parámetros de precisión diagnóstica) (Gonzales, 2012) corresponde a la metodología en el marco de la Investigación Clínica, la cual se realiza en el entorno clínico, dispuesto a dar resolución con alto rigor científico de preguntas relacionadas a la etiología, pronóstico, tratamiento o como en el caso de la presente, al diagnóstico psicológico en contextos clínicos, con la intención de manejar desde el punto de vista epidemiológico las tres variables intervinientes para encontrar asociaciones explicativas o comparativas respecto a la formulación del diagnóstico, a saber, el azar, el sesgo o la mera confusión (García, 2012).

ESTUDIOS DE PRECISIÓN DIAGNÓSTICA:

En el contexto clínico, la utilización de pruebas casi siempre adquiere una orientación diagnóstica, por lo que muchas de las pruebas clínicas también son denominadas diagnósticas, las cuales son de vital importancia en la toma de decisiones para la múltiples series de detección, es decir, pruebas bioquímicas, moleculares o genéticas, así como los signos o síntomas que caracterizan conjuntos clínicos posibles de estudiar y tratar, por lo que se llevan a cabo para reducir la incertidumbre sobre la patología de interés (Rodríguez, 2013), las pruebas diagnósticas son por tanto, cualquier proceso más o menos complejo, orientado a determinar la presencia (detección) o ausencia (descarte) de alguna condición, probablemente patológica, enfermedad o factor de riesgo, no sensible a ser descrita por observación directa

(sino por la traducción clínica de sus manifestaciones; apoyadas por pruebas de cuantificación determinante).

Al hablar de pruebas diagnósticas también se hace referencia a sus elementos, los cuales además de estar compuestos por otras diversas pruebas, también están compuestas por la anamnesis, exploración física, examen mental, análisis de laboratorio, estudios por imágenes, etc. En este sentido, los estudios de validez diagnóstica tienen el fin de garantizar que sólo las pruebas diagnósticas precisas sean utilizadas bajo los mejores niveles de jerarquía que demuestre mejores parámetros de precisión diagnóstica (Rodríguez, 2013; Gonzalo, 2013).

Los estudios de validez diagnóstica son estudios de tipo descriptivo (excepto, cuando se seleccionan dos muestras a partir del diagnóstico, es decir, cuando se determinan casos y controles hablamos de estudios analíticos), por lo que mantienen la finalidad de describir a través del análisis estadístico (fundamentado en la estadística epidemiológica clínica) características de un problema de salud en una población (Argimon y Jiménez, 2013), “se constituyen estudios observacionales ya que corresponden a diseños de investigación clínica en la que se observa y registra fenómenos clínicos de estudio sin intervención o tratamiento (manipulación intencionada) alguno de las variables en el curso natural de estos” (Manterola, 2009, p. 540), teniendo en su mayoría de estudios el objetivo ulterior de identificar asociaciones estadísticamente significativas (las cuales serán ampliadas en estudios analíticos posteriores) de relaciones causales, aspecto no relevante para los estudios de validez diagnóstica, ya que al demostrar una asociación estadísticamente significativa entre el resultado de una prueba y la enfermedad o factor de riesgo no se determinaría la utilidad que tendría la prueba clínica, sino al estimar parámetros descriptivos de precisión/exactitud diagnóstica con sus correspondientes

intervalos de confianza y la comprobación estadística de los mismos (Hulley, Cummings, Browner et al. 2014).

En este sentido, los estudios de validación diagnóstica responden como parte de su metodología problemas de investigación relativos al diagnóstico bajo la arquitectura compuesta por 4 fases (Knottnerus, Van Well, Sackett et al. 2002):

- Fase I: En esta fase el estudio responde la pregunta de si los resultados son distintos para individuos que presentan la enfermedad o factor de riesgo de aquellos que no la presentan.
- Fase II: Por otro lado, responden también si los individuos con determinados resultados de la prueba son más propensos a tener la enfermedad o factor de riesgo que los individuos con otros resultados.
- Fase III: En esta fase se determina si la prueba diagnóstica discrimina a los individuos con y sin la enfermedad, respecto a individuos que clínicamente demuestran mayor probabilidad razonable para sospechar que la enfermedad está presente.
- Fase IV: Finalmente, se responde a la pregunta de que si los pacientes que en los que se administró esta prueba diagnóstica demuestran mejores consecuencias de salud que a pacientes de similar espectro clínico que no fue testado.

Con base a ello, el objetivo central de los estudios que evalúan una prueba diagnóstica son la de estimar su capacidad discriminativa entre personas que padecen alguna enfermedad o que se ven expuestas a factores de riesgo que afectan su estado de salud de las que no cumplen dichos criterios, pero que presentan conjuntos sintomatológicos similares en el mismo rango del espectro clínico (Argimon y Jiménez, 2013). En la que sus desafíos metodológicos se

enmarcan básicamente en la determinación relacional compleja con el gold standard, el manejo de precisión para amplios espectros, el control estadístico de los sesgos, la variabilidad de la muestra apropiada y, sobre todo, la posibilidad de equiparar los avances del desarrollo de tecnologías diagnósticas con el retraso metodológico de investigación para el diagnóstico mismo (Knottnerus, Van Well, Sackett et al. 2002).

Podemos inferir por tanto que la resolución estadística del problema relacionado al diagnóstico puede medirse de forma válida, confiable y reproducible, mediante la inclusión de un procedimiento de referencia o gold standard respecto a la evaluación de utilidad de una nueva prueba, es decir, tras la comparación de una variable predictiva de algún resultado y una variable de desenlace que determina la presencia o no de la enfermedad, la cual es determinada por el procedimiento diagnóstico estándar (Argimon & Jiménez, 2013; Gonzalo, 2013).

Como se ha señalado en la sección del marco referencial y diagnóstico, los estudios de validez diagnóstica son escasos en nuestra especialidad (sobre todo para las variables clínicas de estudio) y los existentes no están precisamente formulados bajo criterios de rigor metodológico de Investigación Clínica, por lo que las pruebas recomendadas por demostrar indicadores de validez diagnóstica son escasas y podrían conllevar parámetros de significancia estadística sujeta a sesgos no controlados por su demarcación metodológica.

Prueba de referencia o gold standard:

Denominada también variable de desenlace (por determinar la presencia o no de la enfermedad); la prueba de referencia, no es más que el procedimiento diagnóstico (prueba diagnóstica) que se define como el “mejor disponible” (énfasis añadido) para discriminar a pacientes que padecen (siempre positivo en pacientes que padecen la enfermedad) y no

(siempre negativo en pacientes que no padecen la enfermedad) la enfermedad o condición de estudio (Argimon y Jiménez, 2013; Burgos y Manterola, 2010), es decir, la que demuestra mayor certeza estadística (prueba más fiable) de detección y descarte en la práctica clínica habitual, pudiendo haber sido determinada como tal debido al respaldo de estudios de validez diagnóstica realizados al momento de su construcción o posteriormente como parte del aporte de consistencia por investigadores, sin embargo, dicha prueba de referencia también puede haber sido definida como tal por mera experiencia o frecuencia de tiempo de uso en la práctica habitual clínica por parte de los profesionales de la salud, interés precisamente a controlar por este tipo de estudios.

Criterios de Verificación:

Respecto a la determinación de la prueba de referencia existen varios aspectos o características a verificar, de entre estos, el patrón de referencia *debe ser el mejor método accesible* para certificar la presencia de la enfermedad, con frecuencia determinada por el estándar que la comunidad científica establece como el más adecuado, por lo que en estudio si la prueba de referencia no demuestra tal impresión es probable que la validez de los resultados disminuya (Rodríguez, 2013), “limitación identificada en la mayoría de los estudios de validación incluidos en la presente como parte del marco referencial por carencia de estudios que precisen parámetros de referencia en nuestra especialidad, más que los basados por consenso de expertos establecidos por la OMS” (Vacca y Mercado, 2005, p. 31); por otro lado, una buena prueba diagnóstica es la que ofrece sencillez en su aplicación, costes mínimos por utilización, alto grado de aceptación por la población clínica por mínima invasión o que generen también mínimos efectos adversos, por lo que sus complicaciones metodológicas que forman parte del problema que fundamente algún estudio de precisión para contrastar su utilidad con la prueba nueva radica en características diametralmente opuestas a las

mencionadas, tratándose por tanto, de elaborar la validez de la nueva prueba diagnóstica determinando su validez de convergencia con un gold standard (Gonzalo, 2013).

En la presente, la *Escala de Depresión de Zung se establece como el patrón de referencia para determinar la presencia o ausencia relacionada a la numeración cuantitativa de severidad para el conjunto clínico*, con base por un lado por la recomendación de la comunidad científica en el contexto de estudio (Ciudad de Huancayo), en la que es la ampliamente utilizada para detectar o descartar ambos conjuntos clínicos y por otro lado gracias a los parámetros de validez diagnóstica determinados por estudios que se incluyeron en el marco referencial.

Sesgos Para su Formulación:

De ente otros aspectos a evaluar respecto al patrón de oro, se encuentra el control ético y metodológico de las no respuestas o casos no estudiados (casos negativos) por el patrón de oro, de tal manera que la prueba de referencia debe administrarse a todos los casos en cuestión, no limitando su aplicación para aquellos a los que se determinó un resultado positivo, controlando de esta forma el *sesgo de verificación* por condicionamiento a los resultados de la prueba nueva, bien porque puede no ser tolerado por lo pacientes o puede ser tan invasivo o costoso que se aplique sólo a los casos más graves. Por otro lado, debemos asegurarnos que la aplicación, los resultados y la interpretación de ambas pruebas sean realizados de forma ciega entre sí [no influenciada por el conocimiento del resultado de la prueba de referencia o viceversa], evitando que su validez se demuestre como artificialmente aumentada o disminuida, controlando *sesgos de subestimación, sobreestimación e incorporación* y en el caso de utilizar un mal criterio de referencia en la que la nueva prueba diagnóstica sea mejores parámetros, la nueva prueba podría resultar no útil, lo mismo para nuevas pruebas muy específicas, produciéndose el *sesgo*

de criterio de referencia imperfecto (Argimon & Jiménez, 2013; Rodríguez, 2013; Burgos et al. 2010).

ESTRATEGIA DE INVESTIGACIÓN:

Los estudios de validez diagnóstica consisten en administrar en una muestra determinada tanto la prueba de referencia como la nueva prueba en estudio, facilitando así con la prueba de referencia la correcta diferenciación de pacientes que padecen y no la enfermedad o condición clínica de estudio, de manera que se pueda evaluar la capacidad discriminativa de la nueva prueba diagnóstica entre ambos grupos (Argimon y Jiménez, 2013), en el marco de los fundamentos de evaluación epidemiológica que se orientan a determinar la mejor agudeza (precisión y exactitud) en la medición (Gonzalo, 2013). Al respecto y, antes de continuar, es oportuno diferenciar conceptos relacionados a la planificación de mediciones cuando hablamos de estudios de validez diagnóstica, por lo que Hulley, Cummings, Browner et al. (2014) en esencia se distinguen la Precisión de la Exactitud:

- **Precisión:** Hace referencia al grado de consistencia, estabilidad o probabilidad de reproducibilidad en que una variable presentará valores cercanos a los de su estimación inicial, siendo posiblemente afectado por errores sistemáticos, es decir, productos del azar no controlados estadísticamente ya sea por parte del observador, los participantes o el instrumento mismo (p. 40).
- **Exactitud:** Se refiere por contrariedad a la precisión al grado en que la variable se aproxima al valor verdadero de medición, por lo que su mejor modo de evaluación sería a través de un patrón de referencia como contrastación para sus parámetros, siendo

posiblemente afectado por errores sistemáticos, es decir, sesgos atribuidos al observador, participantes o la prueba misma (p. 40).

Ambos términos relacionados a los planes de medición no deberían confundir la denominación metodológica del presente estudio, ya que al hablar de estudios de validez diagnóstica hablamos de la utilidad clínica que la nueva prueba puede demostrar con respecto al patrón de oro, en tal sentido, se analizan su exactitud (validez), reproductibilidad (fiabilidad) y viabilidad (confiabilidad) (Hulley, Cummings, Browner, 2014), de modo que al denominar al estudio como uno de “*Precisión diagnóstica*” no se hace más que apelar al consenso que delimita estándares de calidad para la formulación metodológica (con ello su denominación conceptual) de este tipo de estudios, los cuales recomiendan una lista de comprobación para fundamentar la formulación y publicación de estudios como la presente bajo criterios de la Medicina Basada en la Evidencia, en este sentido, con la meta de establecer una herramienta para valorar la calidad de estudios de precisión diagnóstica el presente estudio apela a la declaración STARD (Standards for Reporting of Diagnostic Accuracy) en la cual se debe identificar a estos estudios en términos de *Precisión Diagnóstica* (González y Velarde, 2012), sin inferir en su diferenciación terminológica de la palabra precisión y exactitud.

Continuando con la estrategia metodológica para la investigación de pruebas diagnósticas, y como ya se mencionó que en este tipo de estudios se comparan las pruebas a través del análisis de distribución de una variable (prueba estándar) con relación a otra (nueva prueba), con la finalidad de hallar un patrón relacional significativo o no (a lo que llamamos contingencia), pudiendo a partir de ello y de entre muchas posibilidades que ofrece este procedimiento identificar la probabilidad de que algún grupo padezca alguna enfermedad, en el caso de los estudios que evalúan una prueba diagnóstica al comparar el resultado de la nueva prueba

(predictiva) con la determinación inicial de la prueba de referencia se puede expresar los resultados en distintas formas de acuerdo al tipo de la variable (Cualitativa nominal, cualitativo ordinal, cuantitativo discreto o continuo) (Gonzalo, 2013).

Por lo que en medicina o ciencias de la salud una buena herramienta para expresar este tipo de distribuciones son las tablas de contingencia, específicamente las de doble entrada, en la que se inserten información de dos variables, traducida como frecuencias para cada composición de las variables estudiadas, identificando para el caso las filas al resultado de la variable en estudio (variable 1) y las columnas a determinación de la presencia o ausencia de la enfermedad (variable 2) (Gonzalo, 2013), la validez de estos cálculos se fundamenta en evitar el ingreso de observaciones sesgadas a dos o más valores y en sentido contrario que algunas observaciones no sean ingresadas en algún valor específico.

Tabla 14
Tabla de análisis de distribución de contingencias 2x2.

		Variable 2		Total
		Valor A	Valor B	
Variable 1	Valor X	---	---	---
	Valor Y	---	---	---
Total		---	---	---

Fuente: Gonzalo. (2013). Validez de los elementos diagnósticos en endometriosis (p. 138).

Para la práctica metodológica relacionada a estudios de validez diagnóstica las relaciones que expresan la probabilidad o posibilidad (relaciones matemáticas existentes entre el fenómeno clínico de padecer una enfermedad, condición clínica o estar expuesto a factores de riesgo que afecten la salud) de haber detectado o descartado (grado de certeza o incertidumbre) correctamente la presencia o ausencia más certera de la enfermedad propone 4 tipos de relación posibles, expresados en una tabla de contingencia de 2x2 (también denominada matriz de

decisiones: constituida por el estado real de la condición mórbida y el resultado de la prueba a estudiar), en las que dos probabilidades son correctas y 2 no (Fletcher y Fletcher, 2007), por lo que la comparación de la prueba de estudio y el patrón de referencia expresan en primera instancia la distribución de las relaciones probables, siendo estos 4 tipos de resultados los denominados: Verdaderos Positivos (VP), Verdaderos Negativos (VN), Falsos Positivos (FP) y los Falsos Negativos (FN) (Anza, López, González et al. 2015).

Para evitar sesgos en su formulación ambas pruebas deben ser aplicadas a todos los sujetos, por lo que el patrón de referencia deberá ser aplicado tanto a los que dieron positivo o negativo con la nueva prueba, procurando incluir a pacientes de amplio espectro de severidad (desde el menor grado de severidad hasta el mayor e incluso los que representan retos diagnósticos por comorbilidad, atipicidad, dualidad, etc.), aspectos que se tratarán con mayor profundidad en otra sección.

Tabla 15
Distribución de resultados en un estudio de prueba diagnóstica.

		Enfermedad		Total
		Enfermos (Presente)	Sanos (Ausente)	
Prueba	Positivo (+)	(a) Verdadero Positivo (VP)	(b) Falso Positivo (FP)	<i>Total, de (+)</i> a+b
	Negativo (-)	(c) Falso Negativo (FN)	(d) Verdadero Negativo (VN)	<i>Total, de (-)</i> c+d
Total		<i>Total, de enfermos</i> a+c	<i>Total, de sanos</i> b+d	<i>Total, de individuos</i> a+b+c+d

Fuente: Adaptada de Argimon. (2011); Anza. (2015) & CTO. (2011).

Por lo tanto, la exactitud de una prueba diagnóstica radica en su capacidad discriminativa para clasificar correctamente a los individuos en función de una referencia, ya que no todas las pruebas han sido diseñadas bajo parámetros de certeza diagnóstica (en función a la tasa de desaciertos FP+FN) a diferencia de otras que demuestran mayor precisión (en función a la tasa de aciertos VP+VN); en función a las distribuciones (clasificaciones) encontradas pueden existir los siguientes parámetros que reflejan probabilidades respecto a la enfermedad y la prueba (Argimon & Jiménez, 2013; Hernández, Gil, Delgado et al. 2005):

- Número de pacientes con la enfermedad, diagnosticados como + por la prueba: la cual expresa la tasa de VP, cuando los resultados son positivos en pacientes enfermos.
- Número de pacientes sin la enfermedad, diagnosticados como + por la prueba: la cual expresa la tasa de FP, cuando los resultados son positivos en pacientes sanos.
- Número de pacientes con la enfermedad, diagnosticados como – por la prueba: la cual expresa la tasa de FN, cuando los resultados son negativos en pacientes enfermos.
- Número de pacientes sin la enfermedad, diagnosticados como – por la prueba: la cual expresa la tasa de VN, cuando los resultados son negativos en pacientes sanos.

No es posible continuar con los cálculos que aportan parámetros de validez diagnóstica, sin antes precisar la importancia que tiene el cálculo o determinación de la muestra en este tipo de estudios, puesto que de ello depende también el intervalo de confianza (IC) en los que se expresará los parámetros de precisión diagnóstica (con ello el control de errores aleatorios/productos del azar) la cual es complementaria al error de tipo alfa y la potencia de

un estudio (margen de error) complementaria al error de tipo beta, errores que afectan más directamente a la decisión de rechazar o aceptar nuestras hipótesis.

Errores Alfa y Beta:

El no control de estos dos tipos de errores puede afectar la interpretación adecuada de los datos resultantes del estudio, más exclusivamente en estudios de precisión diagnóstica en la que las hipótesis que se plantean son de tipo estadístico, por la que se establece posibles parámetros resultantes acerca de una población que tendrán que ser rechazados o afirmados (Gonzalo, 2013; Gordillo, Medina & Pierdant, 2012).

Los errores de tipo 1 o errores alfa ocurren cuando se rechaza la hipótesis nula aceptándose la alterna, cuando en realidad no existe diferencia entre las observaciones, es decir, se rechaza la hipótesis nula siendo cierta, dando como significativo algo que no lo es (dependen del IC).

Los errores de tipo 2 o errores beta, son la contraparte del error alfa, por lo que suceden cuando se rechaza la hipótesis alterna y se acepta la nula, cuando la diferencia entre las observaciones si es existente, es decir, no se rechaza la hipótesis nula siendo falsa, dando como no significativo algo que lo es. (Dependen del margen de error).

La conceptualización gráfica se refleja en la siguiente tabla.

Tabla 16

Modelos de error al determinar la muestra y formular hipótesis.

		Su decisión	
		"No rechazar H_0 "	Rechazar H_0
La verdad	H_0 verdadera	Correcto	Error tipo I (riesgo $\alpha-1$)
	H_0 falsa	Error tipo II (riesgo β)	Correcto

Si decide que X es importante (o rechace H_0)
¿A qué clase de error de decisión está expuesto? → Tipo I

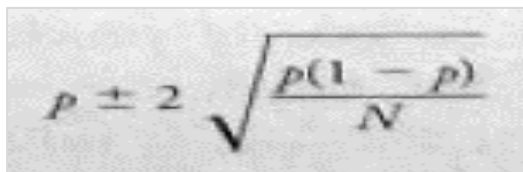
Si decide que X no es importante (o no rechace H_0)
¿A qué clase de error de decisión está expuesto? → Tipo II

Fuente: Gordillo et al. (2012). Manual de Investigación Clínica (p. 41).

Con el Intervalo de Confianza (IC) o también llamado nivel de confianza podemos proporcionar precisión estadística a las estimaciones de indicadores estadísticos que determinan propiedades del elemento diagnóstico, sobre todo en estudios de validez diagnóstica, debido a que los parámetros de validez diagnóstica son estimaciones que describen la población en estudio, la cual se encuentra sujeta a múltiples factores para la variabilidad aleatoria (posibilidad de resultado debido al azar), cuantificando de tal manera este posible error aleatorio y diferenciar entre lo estadísticamente significativo del azar; expresando de esta manera los indicadores estadísticos junto a un intervalo de confianza dentro del cual se encontrará la magnitud real del parámetro de interés no con un grado absoluto, sino con un cierto grado de seguridad, la cual será del 95% para el presente estudio; por lo que de las muestras estudiadas en referencia a la población se espera que en el intervalo del 95% de ellas se hallará el parámetro y en 5% de muestras no, p. ej. para una sensibilidad de 84 (IC 95%: [79;89]), lo cual indicaría que la dicha estimación existe un 95% de probabilidad de hallar el valor real de Sensibilidad entre 79 y 89 (en este sentido, el aumento del tamaño de la muestra mejoraría el IC para los resultados hallados) (Gonzalo, 2013).

$$\text{Intervalo de confianza} = 1 - \alpha (1 - 0,05) = 95\%$$

El cálculo del intervalo de confianza del 95% de una proporción puede realizarse también con la siguiente fórmula (Fletcher & Fletcher, 2007):



$$P \pm 2 \sqrt{\frac{p(1-p)}{N}}$$

Figura 5 Cálculo del IC al 95% para proporciones.

Fuente: Fletcher & Fletcher. (2007). Epidemiología clínica (p.49).

Cuando hablamos de la precisión del estudio o margen de error, nos referimos a la capacidad de detección (agudeza en la medición) de diferencias entre observaciones cuando estas existen en realidad, por lo que se correlaciona a la probabilidad de demostrar la H_1 siendo cierta (capacidad de una prueba estadística para rechazar la H_0 siendo falsa). Su mayor dependencia se consigue con el incremento del tamaño muestral, la cual por consenso equivale al valor de potencia de una prueba estadística del 90%, designando un 10% de probabilidad de error beta para las diferencias o resultados hallados, algunos otros autores establecen un nivel máximo de error beta tolerable del 20% (0,20) por acuerdo de la comunidad científica, la cual conjetura una precisión absoluta de al menos 80%, para la presente investigación se determina un margen de error del 5% (0,05) a modo que esta se integre con el valor del IC (Gonzalo, 2013).

$$\text{Potencia de un estudio} = 1 - \beta (1 - 0,05) = 95\%$$

De esta manera la distribución inicial de resultados no es más que los datos provenientes de un análisis epidemiológico que pretende identificar la prevalencia de alguna enfermedad datos que necesitan ser expresados a través de un Intervalo de Confianza (IC) y Precisión Absoluta

(PA) para otorgar consistencia a los resultados de los que puedan obtener diversos parámetros de precisión que demuestre que la prueba respecto al patrón de oro es válida (Greenhalgh, 2016), a dichos parámetros resultantes de las distintas distribuciones (relaciones que existen entre una prueba diagnóstica en estudio y el patrón de oro) que la tabla de contingencia 2x2 (también denominada matriz de decisiones) resume (Fletcher y Fletcher, 2007) se calculan diferentes características operativas de las pruebas diagnósticas, constituyentes que representan los elementos primarios de las características operativas de las pruebas diagnósticas (Sierra, 2003, p. 181).

El cálculo de los parámetros de precisión o características operativas de la prueba diagnóstica permite obtener diversos índices de utilidad de la prueba (Argimon y Jiménez, 2013), entre ellos la Prevalencia, Sensibilidad, Especificidad, Valores Predictivos Positivos, Valores Predictivos Negativos, Valor Predictivo Global, Precisión Global, Cociente de Verosimilitudes Positivos y Negativos, Índice de Eficiencia Pronostica, Proporción de sanos, Odds pre y post test, Índice Kappa (Sierra, 2003; Donis, 2012) y algunos otros que determinan la precisión, en términos de validez, confiabilidad y fiabilidad de una prueba diagnóstica.

Parámetros Estadísticos de Validez Diagnostica:

Una forma de calcular la validez de una prueba diagnóstica es a través del cálculo de los índices de Sensibilidad (Se) y Especificidad (Sp), de tal manera que se pueda estimar la magnitud de la prueba para cumplir el cometido para el cual fue desarrollado, a saber, discriminar entre personas que padecen una enfermedad de las que no, ya sea en connotaciones dicotómicas (positivo y negativo) o para un espectro de severidad (estudios con resultados continuos: con rangos de severidad; en los que hay que establecer puntos de corte, los cuales explicaremos más adelante), por lo tanto ambas permiten comparar la eficacia de una nueva

prueba con la de referencia o patrón de oro (Henquin, 2013). Estos indicadores demostrarán su potencialidad cuanto más cercano al 100% estén, condición imposible para una prueba de ambos indicadores, de lo que se trata es de equilibrar el mejor desempeño con puntos de corte, los cuales serán vistos más adelante.

La *sensibilidad* responde a la pregunta ¿cuál es la eficacia de la prueba para detectar a las personas con la enfermedad? indica, por lo tanto, la proporción del total de enfermos que el test es capaz de detectar, su respuesta es una expresión de la probabilidad condicional que tiene un individuo enfermo de dar positivo en un test (Greenhalgh, 2016; Gonzalo, 2013; CTO, 2011), se halla de la siguiente manera.

$$Se = VP/(VP+FN) = a/a+c$$

La *especificidad* sería la contraparte al primer indicador de precisión, esta responde a la pregunta ¿cuál es la eficacia de la prueba para excluir correctamente a las personas sin la enfermedad? Indica, por lo tanto, la proporción de sanos confirmados como tales por el test, su respuesta es la expresión de la probabilidad condicional que tiene un individuo sano de dar negativo en un test (Greenhalgh, 2016; Gonzalo, 2013; CTO, 2011), se estima de la siguiente forma.

$$Sp = VN/(VN+FP) = d/b+d$$

Las ventajas que demuestra la sensibilidad y especificidad son que su validez no depende de la prevalencia, así como también es necesario su interdependencia para hallar la validez de la prueba, ya que es imposible una prueba diagnóstica potente para ambas capacidades, de lo que se trata es de calibrar la mejor sensibilidad y especificidad que puede ofrecer una prueba,

además son muy útiles para el cálculo del resto de características operativas, su utilidad es inmediata sobre las decisiones clínicas puesto que se comienza a trazar una jerarquía de precisión en la práctica habitual, sin embargo, no modifican el conocimiento que se tenía sobre el estado del paciente (Gonzalo, 2013); no debemos olvidar que como regla general tanto la sensibilidad como la especificidad informan acerca de la prueba en general (en función del estatus de enfermedad) (Greenhalgh, 2016). Es importante también identificar “la tasa de errores que pudo cometer la prueba al realizar el estudio, de modo tal que se vuelvan también indicadores no de precisión, sino del margen de error con la cuenta cada prueba, los dos más importante se mencionan, como sigue” (González & Velarde, 2012, p. 2):

La *tasa de falsos negativos* la cual expresa el error de haber clasificado a un individuo enfermo como sano, dicha proporción resultante debería ser equivalente al 100% de la muestra en estudio si la sumamos con la proporción esperada por la *sensibilidad*, la TFN se determina de la siguiente manera.

$$TFN = FN/(VP+FN) = c/a+c$$

La *tasa de falsos positivos*, esta refleja el error de haber clasificado a un individuo sano como enfermo, dicha proporción resultante debería ser equivalente al 100% de la muestra en estudio si la sumamos con la proporción esperada por la *especificidad*, la TFP se halla de la siguiente manera.

$$TFP = FP/(VN+FP) = b/b+d$$

Una vez se determine las primeras características operativas, será indispensable hallar la *precisión o valor global de eficiencia*, la *prevalencia de la enfermedad* y la *proporción de sanos*; la precisión responde a la pregunta ¿qué proporción de todas las pruebas han sido el resultado correcto? es decir, indica la proporción de VP y VN respecto al total de resultados, la prevalencia refleja la proporción de individuos de una población que presenta la enfermedad a diferencia de la proporción de sanos que refleja explícitamente su cometido, se hallan de la siguiente manera respectivamente (Greenhalgh, 2016; Gonzalo, 2013; CTO, 2011).

$$VG = \frac{VP+VN}{VP+VN+FP+FN} = \frac{a+d}{a+b+c+d}$$

$$P = \frac{VP+FN}{N} = \frac{a+c}{N} \text{ (Que es la población total en estudio)}$$

$$P_s = \frac{FP+VN}{N} = \frac{b+d}{N}$$

De éstas 3 el *valor global de eficiencia* demuestra la probabilidad de que la prueba diagnóstica acierte en sus conclusiones, de la que se espera una potencia de al menos 80% o más, para poder estimar que la prueba diagnóstica correctamente al 80% de los pacientes con un margen de error del 20% de diagnósticos analizados, sus ventajas se basan en la integración que hace de la sensibilidad y especificidad, además de su interpretación práctica (Gonzalo, 2013).

Parámetros Estadísticos de Confiabilidad Diagnóstica:

Con los primeros parámetros hallados se hace posible reducir considerablemente el grado de incertidumbre respecto a la probabilidad diagnóstica que demostrará alguna prueba tras su administración, por lo que tanto la Se y Sp se consideran características operacionales primarias o fundamentales, sin embargo, estas son limitadas, ya que el profesional requiere también conocer la medida en que sus resultados modifiquen realmente el grado de conocimiento que

se tenía sobre el estado del paciente, es decir, la probabilidad de que en un individuo que se haya obtenido un resultado + efectivamente se encuentre presente la enfermedad o en caso opuesto (probabilidad de que en un individuo que dio negativo por la prueba efectivamente no se encuentre presente la enfermedad), entonces además de preguntarnos cuán discriminativa es una prueba diagnóstica es también necesario preguntarnos de cuán confiables pueden ser aquellos parámetros que ofrece, estas cuestiones se relacionan al problema de credibilidad a la que se enfrenta también toda prueba diagnóstica y, son el cálculo de los Valores Predictivos y otros indicadores de predicción que dan respuesta a estos problemas (Gonzalo, 2013; Henquin, 2013).

Estos parámetros de predicción para la credibilidad de los resultados facilitan la formulación de inferencias respecto al estado de enfermedad con base en los resultados previos que la prueba arroja, por lo que tanto los Valores Predictivos como los Cocientes de Verosimilitud son los indicadores más pertinentes para apoyar tales inferencias (Argimon & Jiménez, 2013).

El *valor predictivo positivo* responde la pregunta de cuál es la probabilidad de que una persona tenga la enfermedad habiendo dado como positivo en una prueba, es decir, indica la probabilidad posprueba de una prueba positiva (proporción real de afectados entre los que dieron + en la nueva prueba) (Gonzalo, 2013 y Greenhalgh, 2016) la cual se puede determinar matemáticamente de la siguiente manera.

$$\text{VPP} = \text{VP}/(\text{VP}+\text{FP}) = a/a+b$$

El *valor predictivo negativo* por otra parte responde a la contraparte del VPP, a saber, cuál es la probabilidad de que una persona no tenga la enfermedad habiendo dado negativo en una

prueba, es decir, indica la posprueba de un resultado negativo (proporción real de sanos entre los que dieron – en la prueba) (Greenhalgh, 2016; Gonzalo, 2013), la cual puede ser determinada matemáticamente de la siguiente forma.

$$VPN = VN/(VN+FN) = d/c+d$$

El tamaño muestral o el tipo de selección muestral influenciará significativamente sus parámetros, ya que, si la prevalencia de una enfermedad se incrementa (mayor tamaño muestral), disminuirá la tasa de falsos positivos y con ello incrementará el VPP, inversamente proporcional a este fenómeno si disminuye la prevalencia (menor tamaño muestral), aumenta el VPN por disminuir la tasa de falsos negativos, es términos sencillos, ambos Valores Predictivos dependen de la prevalencia de la enfermedad en estudio, por lo que la interpretación de los valores de predicción debe realizarse teniendo en cuenta la muestra, así como su extrapolación de resultados debe hacerse a poblaciones con una prevalencia cercana a la hallada en el estudio; la siguiente figura demuestra su influencia.

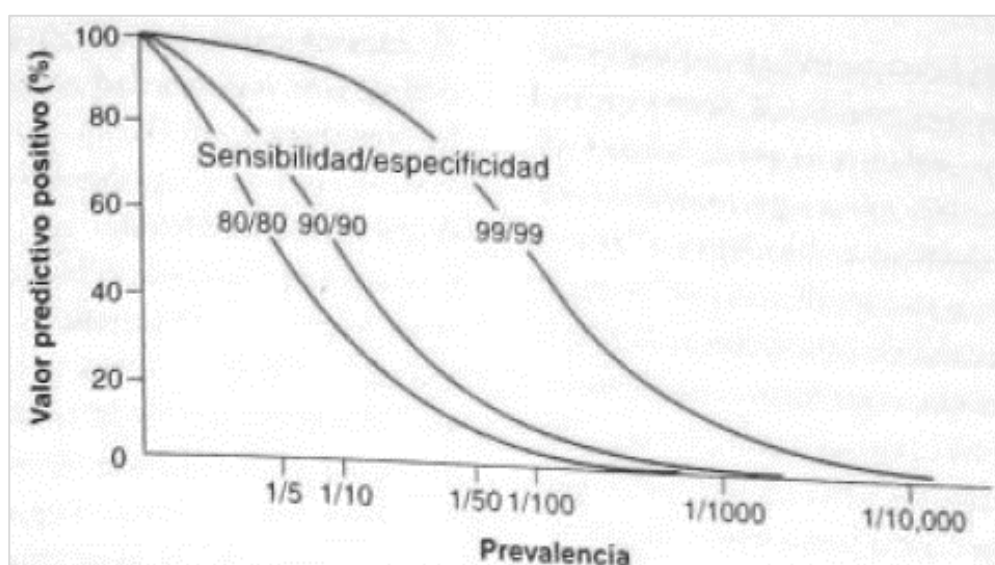


Figura 6 Dependencia muestral para el cálculo de valores predictivos.

Fuente: Fletcher & Fletcher. (2007). Epidemiología clínica (p.51).

Las *ventajas* que demuestran ambos valores de predicción es que si modifican el grado de conocimiento sobre la probabilidad de presentar una enfermedad o condición patológica, además de permitir el cálculo de probabilidad posprueba, su criterio de predictibilidad permite incluir a las pruebas diagnósticas en función a su valor en distintas fases del proceso diagnóstico, sin embargo, se ven afectadas por la prevalencia (Gonzalo, 2013); no debemos olvidar que como regla general tanto el valor predictivo positivo como negativo informan sobre lo que significa un resultado particular de la prueba para un paciente en específico (en función al estatus del resultado del test) (Greenhalgh, 2016).

Ahora bien, hasta el momento hemos descrito características operativas de pruebas diagnósticas que cubren la amplitud de validez y confiabilidad de una medida, sin embargo, una forma alternativa de describir el rendimiento de una prueba diagnóstica es también por medio de la comparación de probabilidades, para lo cual utilizamos la *razón o cociente de probabilidad* también llamado *cociente de verosimilitudes* (denominada en la lengua anglosajona como *Likelihood*=Probabilidad o verosimilitud y *Ratio*=Proporción o razón) con la que comparamos la probabilidad de obtener un determinado resultado en un individuo que presente la enfermedad con la de obtenerlo en un sujeto en el que se ha descartado su presencia (Argimon y Jiménez, 2013; Gonzalo, 2013), su medición se realiza para pruebas positivas y negativas, son de las más complicadas en interpretar y requieren un juicio claro e indemne del investigador o especialista, un estudio de pruebas diagnósticas debe hallar los *cocientes de verosimilitudes positivos y negativos*.

El *cociente de verosimilitud de una prueba positiva* es calculado dividiendo la proporción de casos que tienen un resultado + por la prueba, es decir, la sensibilidad que demuestra, entre la proporción de personas que no tienen la enfermedad habiendo dado un resultado +, es decir,

1 – especificidad, su expresión matemática es la siguiente (Argimon y Jiménez, 2013) la cual permite responder a la pregunta de ¿cuántas veces más probables es que la prueba diagnóstica sea positiva en los enfermos que en los sanos? (Gonzalo, 2013).

$$RV+ = \frac{\frac{VP}{VP+FN}}{\frac{FP}{VN+FN}} = \frac{S}{1-E} ; LR+ = \text{Sensibilidad}/(1-\text{Especificidad})$$

Figura 7 Cálculo matemático de Likelihood Ratio (+).

Fuente: Hernández et al. (2005). Manual de Epidemiología y Salud Pública (p. 82).

El *cociente de verosimilitud de una prueba negativa* es calculado dividiendo la proporción de casos que tienen un resultado –, es decir, 1–sensibilidad, entre la proporción de personas que no tiene la enfermedad haciendo dado un resultado – por la prueba, es decir, especificidad, su expresión matemática es la siguiente (Argimon y Jiménez, 2013) la cual permite responder a la pregunta de ¿cuántas veces más probables es que la prueba diagnóstica sea negativa en los enfermos que en los sanos? (Gonzalo, 2013).

$$RV- = \frac{\frac{FN}{VP+FN}}{\frac{VN}{VN+FN}} = \frac{1-S}{E} ; LR- = 1-\text{Sensibilidad}/(\text{Especificidad})$$

Figura 8 Cálculo matemático de Likelihood Ratio (–).

Fuente: Hernández et al. (2005). Manual de Epidemiología y Salud Pública (p. 82).

A diferencia de los valores predictivos, los Likelihood Ratio reúnen la relación en sus índices con la Sensibilidad y Especificidad, por lo cual no se alteran por el tamaño de la

prevalencia, además de que es útil al emplearla en numerosos niveles de resultados (dicotómicos, continuas, etc.) (Argimon y Jiménez, 2013; Gonzalo, 2013), “*por consenso una buena prueba diagnóstica debe tener un LR- cercano a 0 y un LR+ alta, sin la posibilidad de determinar un límite superior como mejor desempeño*” (Escrig, Martínez & Miralles, 2006, p. 269).

En la literatura más actualizada algunos autores han establecido valores que determinan la eficiencia diagnóstica de los Likelihood Ratio + y – o también denominado índice de eficiencia pronostica (IEP), el cual aporta mayor sustento respecto a los parámetros de eficiencia, se muestran como sigue (Harper & Reeves, 1999; Loong, 2005).

Tabla 17
Capacidad de eficiencia para valores del IEP.

Valores IEP	Capacidad
$IEP(+)\geq 10$ ó $IEP(-)\leq 0.1$	Suficiente
$IEP(+)\geq 5 < 10$ ó $IEP(-) > 0.1 \leq 0.2$	Moderada
$IEP(+)\geq 2 < 5$ ó $IEP(-) > 0.2 \leq 0.5$	Escasa
$IEP(+)\geq 1 < 2$ ó $IEP(-) > 0.5 < 1$	Insignificante

Fuente: Harper & Reeves. (1999); Loong. (2003).

Es posible también determinar el término global del cociente de verosimilitud, a lo que se denomina *razón o cociente de verosimilitud global* para determinar su eficiencia global, sin embargo, esta se aplica para pruebas consecutivas las cuales se calculan básicamente a través de la multiplicación de los valores consecutivos (para elementos diagnósticos consecutivos como entrevista estructurada, anamnesis, prueba, etc.).

Las ventajas del cociente de verosimilitudes son que relacionan la sensibilidad y especificidad en un solo índice, permiten el cálculo de probabilidad posprueba (la cual se verá más adelante al tratar el impacto terapéutico de las pruebas diagnósticas) y pueden además generalizarse a diversas poblaciones a diferencia de los valores predictivos ya que no dependen del tamaño de la prevalencia.

Es preciso determinar que para los cocientes de verosimilitudes equivalentes a 1 indican que el resultado de la prueba no modifica la probabilidad de estar enfermo, si es mayor a 1 si lo haría y si es menor la reduciría, “por lo que sus parámetros de precisión diagnóstica indican que valores >10 (0,1) es una prueba muy potente, de 5-10 (0,1-0,2) que es poco potente, entre 2-5 (0,2-0,5) que los resultados pueden estar viciados y de 1-2 (0,5-1) que la prueba no es útil” (Molina, 2013, p. 171).

Tanto las distribuciones relacionales halladas (nueva prueba y gold standard) como los parámetros que existe determinar un estudio de precisión diagnóstica, se resumen en la siguiente tabla (ver tabla 18).

Tabla 18

Distribuciones y parámetros descriptivos de una prueba diagnóstica.

Estatus de la condición (resultado del gold standard)			
Resultado de la prueba diagnóstica	Presente/positivo	Ausente/negativo	Totales marginales
Test positivo	a (Verdadero positivo)	b (Falso positivo)	a+b
Test negativo	c (Falso negativo)	d (Verdadero negativo)	c+d
Totales marginales	a+c (Pacientes con la enfermedad)	b+d (Pacientes sin la enfermedad)	

Parámetro	Fórmula	Definición
Sensibilidad	$a/(a+c)$	Proporción de pacientes con la enfermedad que tendrán test positivo
Especificidad	$d/(b+d)$	Proporción de pacientes sin la enfermedad que tendrán test negativo
Valor predictivo positivo	$a/(a+b)$	Probabilidad de que el paciente tenga la enfermedad dado que el test es positivo
Valor predictivo negativo	$d/(c+d)$	Probabilidad de que el paciente no tenga la enfermedad dado que el test es negativo
Likelihood ratio (+)	$\text{sensibilidad}/(1-\text{especificidad})$	Describe cuántas veces es más probable que reciba un resultado determinado una persona con la enfermedad que una persona sin la enfermedad
Likelihood ratio (-)	$(1-\text{sensibilidad})/\text{especificidad}$	
Exactitud	$(a+d) / (a+b+c+d)$	La probabilidad de que el resultado del test prediga correctamente la presencia o ausencia de la enfermedad
Odds ratio diagnóstico	$(a/c) / (b/d)$	Razón entre la odds de estar enfermo si la prueba da positivo y la odds de no estar enfermo si la prueba da negativo

Fuente: Bravo & Cruz. (2015). Estudios de exactitud diagnóstica: Herramientas para su interpretación (p. 160).

Parámetros Estadísticos de Fiabilidad Diagnóstica:

Una dimensión de la evaluación de confiabilidad recomendable para la mayoría de estudios clínicos es el acuerdo entre examinadores, exclusivamente “ligada a la capacidad de reproductibilidad de medidas, puesto que en la utilización de pruebas diagnósticas si los resultados que obtenemos son pobremente reproducibles, su valor y utilidad también lo serán”

(Donis, 2012, p. 79), de este modo el control de la reproductibilidad de medidas responde siempre a la pregunta ¿cuán posible es obtener el mismo resultado si repetimos la prueba? Y de entre las diversas fuentes de las que dependen las variaciones en los resultados la *variabilidad intrasujeto* (variación de resultados afectados por factores individuales y por las condiciones de evaluación) y la *variabilidad interobservador* son dos de las más importantes a controlar.

En la presente investigación me evoqué a controlar la variabilidad del segundo tipo, a saber, interobservador, debido tanto a los estándares internacionales de calidad para este tipo de investigaciones, como lo hace una de las directrices establecidas por STARD, para presentar estimaciones relacionadas al control de reproductibilidad de medidas (más específicamente en su punto 24 como parte de su lista de elementos a consideración científica para la publicación) (Altman y Bossuyt, 2005) y, por tanto, por el método más ajustado y pertinente a este tipo de estudios, en la que es preciso estimar estadísticamente la reproductibilidad de medidas como parte del juicio clínico de especialistas; estudios al respecto demuestran la utilidad del control de este tipo de variabilidad más ligado a la concordancia entre juicios clínicos (diagnósticos), como el estudio realizado por Landa, N. y su equipo de investigación en el 2008 en España, en la que determinó la concordancia para la detección diagnóstica en salud mental, en los que a través del cálculo del índice Kappa se estimó un valor global de 0,385 para el diagnóstico de diversos trastornos de salud mental (p. ej. consumo de alcohol, trastornos psicóticos y otros), concluyendo que la concordancia para el diagnóstico entre especialistas es baja (Landa, Goñi, García et al. 2008); indicadores similares al estudio realizado en el 2015 por Pérez, O. y demás, al trazar el objetivo de describir las características de un programa de soporte a la primaria y analizar la concordancia diagnóstica entre médicos de atención primaria y psicólogos, estudiando 769 casos y también siendo analizadas por el estadístico Kappa de Cohen,

encontrando que la concordancia diagnóstica para por ejemplo trastornos de ansiedad y entre otros (depresión=20,2% y trastornos adaptativos=19,6%) fue baja (21%), concluyendo en que la baja concordancia implicaría la necesidad de incrementar espacios de interconsulta, para reducir el margen de error alto en la elaboración de diagnósticos (Pérez, García, Grases et al. 2015).

Como ya se mencionó parcialmente como parte de los estudios presentados, si deseamos controlar la variabilidad entre observadores (especialistas) podemos utilizar la prueba estadística denominada Índice de Kappa o Índice de Concordancia, con esta podemos comparar resultados registrados por dos o más especialistas, con la intencionalidad de estimar la probabilidad con la que estos especialistas hayan concordado sólo por azar, siendo lo relevante, el demostrar que dicha concordancia o acuerdo entre examinadores no ocurrió meramente por cuestiones de azar, sino que se debe al método diagnóstico independientemente del especialista (Henquin, 2013).

El método es relativamente sencillo ya que sólo requiere de una sola sesión de administración de la prueba, consiste básicamente en suministrar la prueba diagnóstica a la muestra en estudio y devolverla a uno o un conjunto de especialistas que los puntuarán independientemente a la prueba o a la valoración de otro especialista, posteriormente se verifica el grado de acuerdo entre jueces (Tornimbeni, Pérez y Olaz, 2008), la cual se calcula de la siguiente manera.

$$\text{Kappa} = \frac{(\text{porcentaje de concordancia observada}) - (\text{porcentaje de concordancia esperada solo por azar})}{100\% - (\text{porcentaje de concordancia esperada sólo por azar})}$$

Figura 9 Cálculo del Índice Kappa.

Fuente: Henquin. (2013). Epidemiología y estadística para principiantes (p. 219).

$Kappa = Fc - Fa / N - Fa$; donde: Fc= % de concordancia observada, Fa= % de concordancia esperada sólo por azar y N= la muestra/100%.

Su determinación equivalente a 1 expresa la máxima concordancia posible, una equivalente a 0 que la concordancia observada se corresponde exclusivamente al azar y una de -1 en que la mayoría de las observaciones discrepan más de lo esperado por el azar (CTO, 2011), también se estiman valores de calidad para el Índice de Kappa los cuales determinan la interpretación de nuestros cálculos (Tornimbeni, Pérez & Olaz, 2008), se muestran en los siguientes.

Tabla 19
Valoración de calidad del Índice Kappa.

Valor de k	Fuerza de la concordancia
< 0,20	Pobre
0,21 – 0,40	Débil
0,41 – 0,60	Moderada
0,61 – 0,80	Buena
0,81 – 1,00	Muy buena

Fuente: Tornimbeni et al. (2008). Introducción a la psicometría (p. 75).

Curvas ROC:

Las Curvas ROC también denominadas por sus siglas en lengua anglosajona Receiver Operating Characteristics, traducida por tanto como *curva de Características Operativas del Receptor*, fueron desarrolladas en el contexto bélico a inicios de la década de los 50, desarrollada por ingenieros eléctricos e ingenieros de radar y más específicamente durante la segunda guerra mundial, su finalidad de uso estuvo orientada a la detección de señales electrónicas y detección de ruidos para determinar la eficacia comparativa de radares (Teoría de detección de señales), a través del análisis de capacidad de detección y predicción de los radares con los que contaban, se analizó uno de los operadores de receptor, inicialmente la

resultante del análisis fue dicotómica, en el sentido de que el receptor podía ser muy eficaz para detectar entre “señales verdaderas” y “señales de otro tipo” en función a su calibración (Burgueño, García & González, 1995).

La metodología mecánica que fundamentó los principios matemáticos de cálculo estadístico para la actualidad se basaban en la calibración del umbral del receptor de detección de señales, el cual permitía registrar menores tasas de error, por lo que a medida que el umbral disminuía la tasa de falsos negativos decrecía (primeros conceptos de sensibilidad) y viceversa, es decir, si el umbral aumentaba la tasa de falsos positivos también decrecía (Gonzalo, 2013; Domínguez y González, 2002, p. 172), en la figura 10 se muestra la gráfica del análisis que se realizó al receptor del radar, en ella se distribuyen múltiples señales analizadas en el espectro de ruido de fondo, en la que la amplitud de las señales varía según su calibración, siendo la I, II y III correspondiente a misiles (Burgueño, García & González, 1995).

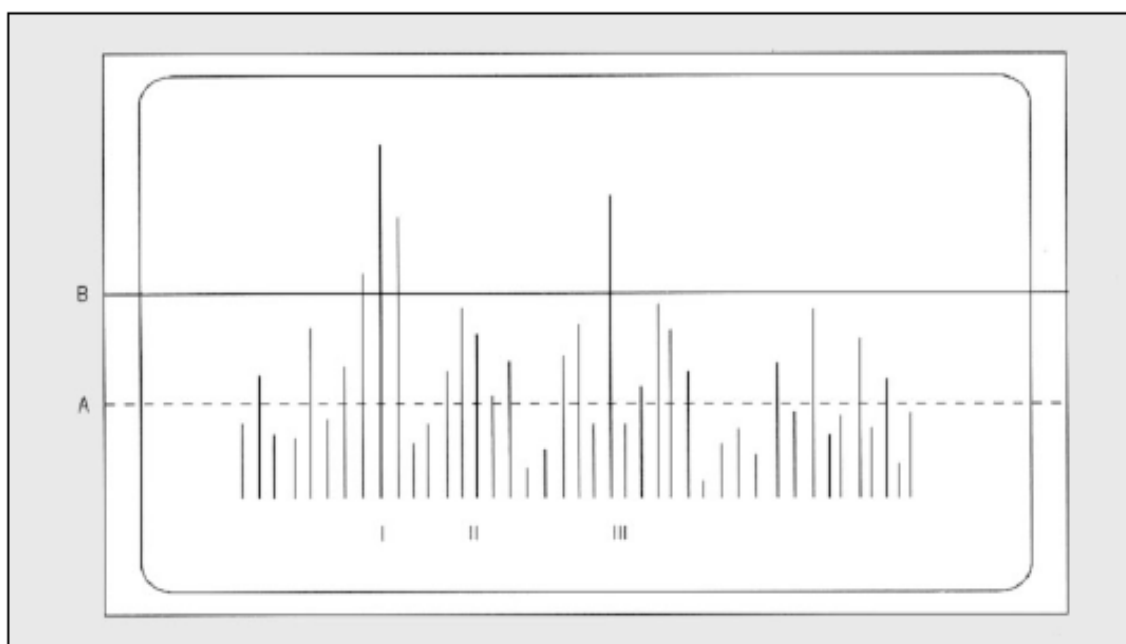


Figura 10 Análisis de operador del receptor.

Fuente: Burgueño et al. (1995). Las curvas ROC en la evaluación de las pruebas diagnósticas (p. 664).

La utilización de las Curvar ROC fue con el paso del tiempo extrapolando sus aplicaciones, siendo utilizado para el control de calidad de sistemas, para el diagnóstico en salud, especialmente en medicina y también en la psicología, contando con un importante representante para este último campo del conocimiento, a saber, el Psicólogo americano John A. Swets quien fue el primero en adaptar la Teoría de la detección de señales al campo de la psicofísica experimental (Green y Swets, 1966), la cual en el año 1967 se convirtió en una herramienta médica de análisis de diagnósticos, siendo Leo Lusted el primero médico (radiólogo) en utilizarla en la toma de decisiones médicas, utilizándolas en estudios con imágenes (Domínguez & González, 2002).

Este análisis fue integrado posteriormente a todo el campo de conocimiento relacionado a la evaluación de pruebas diagnósticas (en todo lo que concierne al diagnóstico en ciencias de la salud), tratándose bajo estos contextos en detectar tras la calibración los mejores puntos de corte (equivalente al receptor del radar) que determinen parámetros operativos para la discriminación diagnóstica de sanos y enfermos, por lo que al ajustar los puntos de corte se controla las tasas de Se y Sp, siendo las curvas una representación gráfica de ambos indicadores primarios, por tanto, también de los demás indicadores de validez diagnóstica, la cual se constituye por dos ejes, en el eje de ordenadas se sitúa los parámetros de sensibilidad y en el eje de abscisas el parámetro complementario de la especificidad ($1 - \text{especificidad}$) (Argimon y Jiménez, 2013; Gonzalo, 2013), en la figura 11 se representa un ejemplo de curva ROC en la que cada punto de la curva señala un valor específico de cociente de verosimilitud positiva equivalente también a un punto de corte específico.

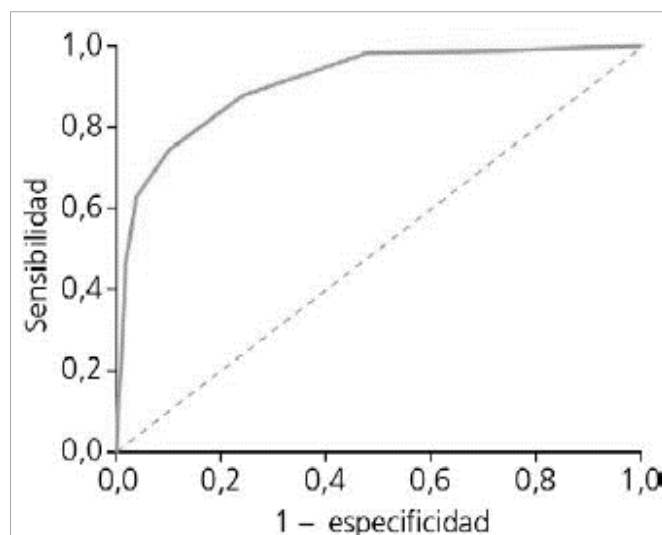


Figura 11 Curva ROC.

Fuente: Argimon & Jiménez. (2013). Métodos de investigación clínica (p. 342).

Las Curvas ROC aplicadas al campo de la medicina o a la evaluación clínica, demuestran índices de exactitud diagnóstica, proporcionando en conclusión un criterio unificador en el proceso de evaluación de una prueba, gracias a sus diversas aplicaciones, de entre las más utilizadas se encuentran:

Tabla 20

Aplicaciones de las Curvas ROC para pruebas diagnósticas.

<p>Evaluación visual de la exactitud Comparación visual entre pruebas Análisis estadístico ROC Evaluación cuantitativa de la exactitud mediante el área bajo la curva ROC Comparación cuantitativa entre pruebas Selección de niveles de decisión considerando la prevalencia y la relación coste/beneficio de los resultados falsos Análisis de regresión logística y análisis discriminante</p>

Fuente: Burgueño et al. (1995). Las curvas ROC en la evaluación de las pruebas diagnósticas (p. 664).

Actualmente en nuestra especialidad se realizan pocos estudios (por ser considerado, puesto que en nuestra localidad el avance de la Psicología no se encuentra muy bien fundamentada por estudios psicológicos con alto rigor científico) al respecto, la inclusión de las Curvas ROC para el análisis de las características operativas de alguna prueba psicológica ni si quiera es una alternativa, en tal sentido, el 2017 Navarro, J. et al. realizaron una revisión sistemática para el

estudio de la depresión y trastornos asociados en adolescentes en el Perú, recopilando 31 artículos, 10 de los cuales fueron descartados antes del análisis principal por no sujetarse a ciertos criterios de inclusión (incluir variables como depresión y muestra adolescente) designados por el estudio, concluyendo tras el análisis (de entre los muchos estadísticos que analizaron las curvas ROC no fueron de las principales, estimando que sólo 3 de todas las publicaciones mostraban sus puntos de corte) que en los estudios destaca la inconsistencia en el reporte y sustento bibliográfico del diseño de investigación, del procedimiento de muestreo y de las propiedades psicométricas de los instrumentos, también de la utilización de instrumentos que no han sido evaluados en nuestro contexto y de punto de corte obtenidos en muestras de otros países, pudiendo afirmar que los estudios no cumplen en su totalidad con los estándares de publicación propuestos por la APA (Navarro, Moscoso y Calderón, 2017); demostrando la infrautilización en las investigaciones, a pesar de ser herramientas fundamentales en la evaluación de pruebas diagnósticas, específicamente en el campo de la medicina o de las prácticas relacionadas a la salud, las cuales se deban quizá a sus supuestas desventajas, mencionadas a continuación (Zweig & Campbell, 1993, p. 568):

- Los puntos de corte no aparecen en su representación gráfica, pese a que al observar una inflexión de la curva se observa la asociación entre la Se y Sp.
- La representación gráfica tampoco muestra el número de sujetos de la muestra.
- Al disminuir el tamaño muestral la curva ROC tiende a la desigualdad.
- Por la dificultad de cálculo de sus parámetros y por la restringida distribución de los programas estadísticos o epidemiológicos.

Punto de Corte:

La práctica clínica no suele ser tan análoga a la realidad diagnóstica bajo parámetros dicotómicos, por lo que las pruebas con resultados que dan positivo y negativo representan hoy en día una menor proporción respecto a las herramientas diagnósticas de uso continuo, siendo las de escalas ordinales o continuas las más utilizadas por su versatilidad para comprender el estado espectral de la salud en la población afectada, siendo necesario a la vez más que sólo 2 categorías para enmarcar la condición clínica en estudio (Gonzalo, 2013). Tratándose, por tanto, la presente investigación de un estudio de pruebas diagnósticas con escalas ordinales que expresan sus resultados en rangos equivalentes a más de 2 puntos posibles para dividir a pacientes probablemente enfermos de los pacientes probablemente no enfermos, se hace hincapié en lo que implica los puntos de corte para un análisis con Curvas ROC.

La elección del punto de corte para la mayoría de estudios actuales en los que se utiliza las Curvas ROC se realizan bajo criterios más flexibles, por lo que determinar un único punto de corte que demuestre la “mejor” sensibilidad y especificidad de la prueba sería proceder de manera estricta (la noción de punto de corte óptimo no es estricta, ya que esta se puede ajustar a las necesidades del especialista), aspecto poco relacionado a la práctica clínica diaria; por lo general cuando analizamos las características operativas en escalas ordinales los mejores puntos corresponden a los puntos de inflexión de la curva (demuestran los límites discriminativos de categorías de los valores de cada rango), posteriormente debemos calcular los cocientes de verosimilitud para cada categoría diagnóstica (Gonzalo, 2013).

Y por recomendación metodológica de los estudios de validez diagnóstica más actualizados se realizarán en la presente *múltiples puntos de corte* para las variables cualitativas ordinales, calculando la Se y Sp de las diversas categorías de las pruebas en estudio y luego elegir los

puntos que optimicen dichos indicadores de discriminación diagnóstica como los sugiere Gonzalo, F. en su tesis el 2013.

Construcción Paramétrica y no Paramétrica:

Dependerá en exclusiva de los tipos de datos que aporte la prueba diagnóstica en estudio, ya sean discretos o continuas (equivalente dentro del rango no paramétrico a la escala ordinal), por lo que se empieza en todos los casos por calcular puntos de sensibilidad y especificidad ($1 - Sp$) a todos los posibles puntos de corte de la prueba, si se trata de una escala continua los puntos serán equivalentes a todos los valores observados a diferencia de la escala discreta en la que los puntos de corte equivalen a todos los posibles valores de la escala, posteriormente el modelo de elaboración de la Curva ROC dependerá de la escala de medida para sujetarse a un modelo no paramétrico (escala continua; por ser no paramétricas también es posible utilizarlas para variables ordinales) o paramétrico (escala discreta) (Burgueño, García y González, 1995); la ventaja central de las primeras radican en la posibilidad de usar todos los datos, pasando la curva por todos los puntos de corte posible y, el segundo modelo ofrece la ventaja de una línea de visualización más clara (Gonzalo, 2013); entre demás ventajas y desventajas que ofrecen ambos modelos se muestran en las siguientes tablas; cabe resaltar que cuando se obtuvieron resultados por la administración de múltiples pruebas, estas pueden ser representadas juntas (curvas superpuestas), las cuales también sirven para comparaciones en las que la posición relativa de las curvas indica la precisión relativa de las pruebas (Domínguez & González, 2002).

Tabla 21

Ventajas e inconvenientes para Curvas ROC no paramétricas.

<p><i>Ventajas</i> Usan todos los datos y la curva pasa por todos los puntos de corte Proporcionan estimadores centrados (sin sesgo) de la sensibilidad, la especificidad y el área bajo la curva, ya que para cada punto de corte los mejores estimadores son los no paramétricos No necesitan asumir supuestos Los cálculos son simples</p> <p><i>Inconvenientes</i> Tienen apariencia de escalera Las series grandes de datos producen un gráfico abigarrado Los empates pueden causar problemas, ya que normalmente dan lugar a curvas subestimadas y, por tanto, la exactitud obtenida es menor que la real La comparación entre dos puntos de dos curvas ROC sólo puede realizarse a sensibilidades o especificidades observadas</p>
--

Fuente: Burgueño et al. (1995). Las curvas ROC en la evaluación de las pruebas diagnósticas (p. 666).

Tabla 22

Ventajas e inconvenientes para Curvas ROC paramétricas.

<p><i>Ventajas</i> Proporcionan una línea curva de visualización más clara. Permiten comparar puntos de dos curvas a cualquier sensibilidad o especificidad, observadas o no</p> <p><i>Inconvenientes</i> Descartan datos al agruparlos, con la consiguiente pérdida de información Asumen modelos de distribución que muchas veces no se pueden comprobar La curva no pasa necesariamente por los verdaderos puntos Los puntos ROC y el área pueden ser sesgados Los cálculos son complejos</p>
--

Fuente: Burgueño et al. (1995). Las curvas ROC en la evaluación de las pruebas diagnósticas (p. 666).

Area Under the Curve (AUC):

El área bajo la curva (ABC) o más conocida en la literatura académica en la lengua anglosajona como area under the curve (AUC), se define como la medida global que refleja que tan buena es la prueba diagnóstica para discriminar pacientes (exactitud) con o sin la enfermedad a lo largo de todo el rango de puntos de corte posibles (Cerdeja y Cifuentes, 2012), representa la probabilidad de clasificar con precisión entre un par de individuos [con y sin la condición patológica], seleccionados por azar (Argimon y Jiménez, 2013). Su expresión sería

$$ABC_{ROC} = \text{Prob} [X_{\text{enfermo}} > X_{\text{sano}}].$$

Esta medida es única e independiente de la prevalencia de la enfermedad, es posible entender este parámetro teniendo en cuenta que el eje Y corresponde a la proporción de VP/Se y que el eje X equivale a la proporción de FP/1-Sp las cuales en su análisis se comparan para cada punto de corte que ofrece una prueba diagnóstica cuya escala de medición es continua (Cerdeja y Cifuentes, 2012), entonces en conclusión el AUC bajo la curva ROC es un indicador global de precisión diagnóstica, la cual hace posible expresar el desempeño de una prueba mediante un valor numérico simple (Domínguez & González, 2002).

En condiciones extremadamente favorables un gráfico en el que la prueba discrimine con perfección entre enfermos y sanos se caracterizaría por una línea ajustada a los dos lados izquierdo y derecho superior del cuadrado (condición imposible) (Gonzalo, 2013), por otro lado, la línea diagonal trazada entre el punto 0,0 y 1,1 se denomina diagonal de referencia o línea de no-discriminación (0,5), describiendo la Curva ROC de una prueba diagnóstica incapaz de discriminar pacientes enfermos de sanos más que por el simple azar (Cerdeja y Cifuentes, 2012), por lo que mientras más alejada esté el trazo de la curva ROC hacia el vértice superior izquierdo (más lejana de la línea de no-discriminación) demostrará mejor capacidad discriminativa (Gonzalo, 2013).

Cálculo del AUC y Valores de Comprobación:

Entonces, una manera de estimar cuantitativamente (determinación de capacidad global) el desempeño discriminativo de pruebas diagnósticas es por medio del cálculo del área debajo de la curva ROC (la cual es útil como índice para comparar la capacidad de pruebas diagnósticas) (Gonzalo, 2013); la cual para una prueba con Se y Sp del 100%, el AUC bajo la curva equivaldría a 1 y cuando la prueba no tenga precisión de discriminación el AUC equivaldría al trazo diagonal con valor de 0,5 (Argimon y Jiménez, 2013). Por lo tanto y por consenso, el

AUC será siempre mayor o igual que 0,5 moviéndose el rango entre la discriminación perfecta (1) y la no existencia de diferencias en la distribución de resultados de la prueba entre los grupos de enfermos y sanos (0,5) (Domínguez y González, 2002); si el caso fuera en el que la curva ROC se sitúe por debajo de la línea de no-discriminación (menos a 0,5) indicaría que la prueba diagnóstica estimó resultados viciados (Gonzalo, 2013).

Los valores de comprobación han sido interpretados por Swets bajo los criterios en los que considera que valores entre 0,5-0,7 indican baja exactitud; si oscila entre 0,7-0,9 pueden estimarse como útiles de acuerdo al propósito o necesidad del investigador, pero en general indican una exactitud moderada y si se estima un valor mayor al 0,9 indicaría exactitud alta o lo que es lo mismo decir, exacta (Argimon y Jiménez, 2013; Burgueño et al. 1995; Swets, 1988); entre las muchas maneras de evaluar la capacidad discriminativa en pruebas diagnósticas es estimando el intervalo de confianza del área bajo la curva, en la que si el valor hallado es superior a 0,5 se considera que la prueba es capaz de discriminar entre grupos de pacientes enfermos de los sanos, esta se puede obtener mediante una prueba de hipótesis o mediante la estimación del IC para AUC, rechazamos la hipótesis si el AUC teórica equivalente a 0,5 [$p < 0,05$], también nos permitiría comparar varias AUC entre múltiples pruebas diagnósticas; por otro lado, el cálculo del error estándar se estimará en relación al tipo de curva producto del tipo de modelo de análisis ya sea paramétrico o no paramétrico (Gonzalo, 2013; Burgueño et al. 1995; Swets, 1988).

Impacto terapéutico de las pruebas diagnósticas:

El impacto que ejerce sobre la toma de decisiones el aplicar o no una prueba diagnóstica se estrecha rápidamente cuando hablamos de la ulterior formulación del plan terapéutico e incluso pronóstico, ya que, si la exploración inicial implica identificar y ordenar la sintomatología del

paciente en cuadros clínicos resultantes y específicos para un trastorno X, la administración de pruebas diagnósticas implicarían ser el vehículo por el cual en función a las necesidades del especialista se puedan cuantificar nuestras observaciones y en medida de la capacidad discriminativa que ofrezca cada parámetro de validez de la prueba se pueda tomar decisiones basadas en evidencia para la posterior formulación del plan terapéutico y la decisión de integrar o no un proceso terapéutico, con lo cual describiríamos también posible predicción pronóstica basado no sólo en características epidemiológicas o clínicos sino también en parámetros de precisión diagnóstica (Zamora, 2008, p. 5), este impacto sobre la toma de decisiones se puede graficar como sigue.

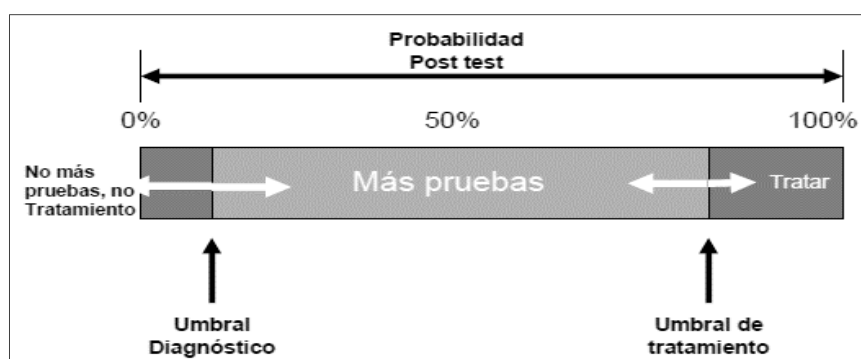


Figura 12 Impacto de parámetro diagnóstico sobre decisión terapéutica.

Fuente: Zamora. (2008). Metodología de evaluación de pruebas diagnósticas (p. 7).

De la figura presentada anteriormente se reflexiona que de entre los muchos aspectos de la evaluación diagnóstica los relacionados al impacto clínico y particularmente a los cambios en el manejo clínico o los cambios en los resultados clínicos de los cuales dependen también los parámetros de exactitud diagnóstica, puesto que la utilización del mismo no se reduce sólo a una única administración, en la mayoría de casos también es utilizada como herramienta de monitorización o seguimiento de la evolución del paciente, por lo que una prueba buena capacidad discriminativa y cocientes de verosimilitud significativas pueden orientar a una probabilidad posprueba también significativa, lo que se resume en condiciones más favorables para la atención del paciente, en tal sentido podemos decir que el enlace entre diagnóstico y

tratamiento es directo y que uno depende del otro, en esencia, el diagnóstico guía todo el proceso posterior a la exposición al paciente, un impacto positivo se deriva de una prueba con significativa Sensibilidad para luego plantear un tratamiento que también demostró significativa eficacia para el problema de salud en específico, además de que se evitan daños por tratamientos inadecuados o innecesarios cuando evitamos Falsos Positivos gracias a pruebas altamente sensibles y por otro lado, a menos precisión diagnóstica, menor calidad de interpretación de la pruebas, por ende, menor calidad para la toma de decisiones terapéuticas, lo que se resume en un procedimiento clínico ineficaz (Zamora, 2008; Otero & Sierra, 2003).

Odds Ratio:

Los cocientes de verosimilitudes sirven para calcular la probabilidad posprueba conociendo la probabilidad preprueba, es decir, la prevalencia y el LR de la prueba, por lo que para hallarla debemos primero expresar la prevalencia en *Odds* (razón que representa la probabilidad de que ocurra un suceso frente a la probabilidad de que no ocurra), siendo los Odds calculados de la siguiente manera: $Odds = p/1-p$: para luego multiplicar la odds preprueba por el cociente de verosimilitud (LR) obteniendo de esa manera la Odds posprueba ($Odds\ posprueba = Odds\ preprueba \times LR$) y finalmente convertimos la Odds posprueba en la probabilidad posprueba ($Probabilidad\ posprueba = Odds\ posprueba / (Odds\ posprueba + 1)$); en este sentido una prueba diagnóstica útil sobre el impacto clínico se caracteriza por generar cambios desde el cálculo de la probabilidad preprueba hasta el de la posprueba; por tanto, si una prueba demuestra un LR+ superior a 10 o una LR- inferior a 0,1 tienen una gran utilidad clínica (Argimon & Jiménez, 2013; Gonzalo, 2013).

Predicción Clínica:

Otra manera relativamente más sencilla de calcular la probabilidad posprueba es a través del análisis gracias al nomograma de Fagan, valiéndonos básicamente de los cocientes de verosimilitudes para calcular la probabilidad posprueba, aunque el método matemático aporte mayor precisión (visto con antelación en el cálculo de Odds para la probabilidad posprueba), sin embargo, si utilizamos las razones de verosimilitud el procedimiento es sencillo, el nomograma de Fagan está constituido por 3 líneas trazadas a partir de la probabilidad preprueba (primera línea), las trayectorias demarcadas por los cocientes de verosimilitudes (segunda línea) las cuales al trazarlas teniendo como punto inicial de ubicación a la línea correspondiente a la prevalencia en un grado determinado por el cociente de verosimilitud hallado, determinando en la tercera línea paralela la probabilidad posprueba de dicha probabilidad (Fletcher & Fletcher, 2007; Greenhalgh, 2016).

NIVELES DE EVIDENCIA EN PROCESOS DE EVALUACIÓN:

Como última píldora, la presente investigación traza la jerarquía de la misma bajo niveles de evidencia y grados de recomendación pertinentes para la formulación metodológica que fundamente un proceso de investigación con alto rigor científico tanto para la actual presentación del proyecto como para la presentación final del informe y con ello su futura publicación, la cual además de ser de menester académico, también es parte de la disposición personal por compartir hallazgos científicos para la comunidad académica en formación y la que ejerce nuestra noble profesión.

En tal sentido, la presente investigación se ha diseñado bajo la lista de comprobación para estudios sobre precisión de pruebas diagnósticas (STARD: Standards for Reporting of Diagnostic Accuracy) (González y Velarde, 2012), la cual también guiará la calidad del

procedimiento científico en recolección de datos, procesamiento y demás pasos para cuando sea aceptado y ejecutado, por otro lado, se tiene previsto evaluar la calidad del mismo para cuando se elabore el informe final de investigación a través de la lista guía de comprobación para estudios de este tipo en el campo clínico (QUADAS: Quality Assessment Diagnostic Accuracy Studies), con la finalidad de aportar mayor calidad al producto de esta investigación, a saber, el artículo científico (González y Velarde, 2012). Las diferencias que ofrecen ambas declaraciones de calidad científica se resumen en la siguiente tabla:

Tabla 23

Diferencias entre declaración STARD y QUADAS.

STARD	QUADAS-2
Para valorar la precisión de un estudio de pruebas diagnósticas.	Para evaluar la calidad de una prueba diagnóstica.
Aspecto a evaluar: la precisión, es decir la capacidad de una prueba para identificar a los pacientes con la enfermedad.	Aspecto a evaluar: el rigor de los estudios primarios.
Su uso es prospectivo: <ul style="list-style-type: none"> • Para los investigadores en la fase de diseño del estudio. • Para los editores. 	Su uso es retrospectivo: <ul style="list-style-type: none"> • Para las revisiones sistemáticas. • Para los metaanálisis.
Evalúa principalmente la validez interna: sesgos y variaciones.	Evalúa principalmente su aplicabilidad.

Fuente: González & Velarde. (2012). Lista de comprobación STARD (p. 4).

Respecto a los niveles de evidencia para el diseño y construcción del proyecto se tomó la referencia a la clasificación que realizó el epidemiólogo David Sackett para jerarquizar los estudios por su nivel de evidencia, siendo 1 correspondiente a estudios que demuestran la “mejor evidencia” y en extremo opuesto a la “peor evidencia” equivalente al nivel 5, de esta manera los diseños del estudio se clasifican tanto por sus grados de recomendación GR (en el grado A las conclusiones se generan por la evidencia más fuerte del estudio; en el grado B las conclusiones son sólo orientativas; para el grado C son menos fiables y de la D no se sugiere tener en consideración referencial) y sus niveles de evidencia NE.

En función a esta clasificación de estudios dirigidas al estudio de un fenómeno relacionado al diagnóstico la presente investigación se encuentra en un NE 1b con grado de recomendación A, se muestra como sigue.

Tabla 24
Clasificación de NE y GR según Sackett.

GR	NE	Terapia, prevención, etiología y daño	Pronóstico	Diagnóstico	Estudios económicos
A	1a	RS de EC con AA	RS con homogeneidad y Meta-análisis de estudios de cohortes concurrentes	RS de estudios de diagnóstico nivel 1	RS de estudios económicos nivel 1
	1b	EC con AA e intervalo de confianza estrecho	Estudio individual de cohortes concurrente con seguimiento superior del 80% de la cohorte	Comparación independiente y enmascarada de un espectro de pacientes consecutivos, sometidos a la prueba diagnóstica y al estándar de referencia	Análisis que compara los desenlaces posibles contra una medida de costos. Incluye un análisis de sensibilidad
B	2a	RS de estudios de cohortes	RS de estudios de cohortes históricas	RS de estudios de diagnósticos de nivel mayor que 1	RS de estudios económicos de nivel mayor que 1
	2b	Estudios de cohortes individuales. EC de baja calidad	Estudio individual de cohortes históricas	Comparación independiente y enmascarada de pacientes no consecutivos, sometidos a la prueba diagnóstica y al estándar de referencia	Comparación de un número limitado de desenlaces contra una medida de costo. Incluye análisis de sensibilidad
	3a	RS con homogeneidad de estudios de casos y controles			
	3b	Estudio de casos y controles individuales		Estudios no consecutivos o carentes de un estándar de referencia	Análisis sin una medida exacta de costo, con análisis de sensibilidad
C	4	Series de casos. Estudios de cohortes y de casos y controles de mala calidad	Series de casos. Estudios de cohortes de mala calidad	Estudios de casos y controles sin la aplicación de un estándar de referencia	Estudio sin análisis de sensibilidad
D	5	Opinión de expertos sin evaluación crítica explícita, o basada en fisiología, o en investigación teórica	Opinión de expertos sin evaluación crítica explícita, o basada en fisiología, o en investigación teórica	Opinión de expertos sin evaluación crítica explícita, o basada en fisiología, o en investigación teórica	Opinión de expertos sin evaluación crítica explícita, o basada en investigación económica

AA: Asignación aleatoria.

Fuente: Manterola, Asenjo & Otzen. (2014). Jerarquización de la evidencia. Niveles de evidencia y grados de recomendación de uso actual (p. 707).

Aspecto relativamente diferente a la clasificación realizada por la Universidad de Oxford (OCEBM), en la que de acuerdo a su consideración la presente investigación se mantendría con un grado de recomendación A, pero en un nivel de evidencia 1 no estable, en el sentido de que la presente sólo estudiará estadísticamente los parámetros de precisión diagnóstica bajo una muestra y no cohortes diferenciadas como lo representa la selección de muestra por casos

y controles (Manterola y Zavano, 2009), debido no a la imposibilidad de este tipo de selección de muestra sino más bien a las ventajas que ofrece estudiar una sola muestra, las cuales se resumen en básicamente en el control de un amplio espectro de pacientes similar al de la práctica clínica habitual, permitiendo mayor posibilidad para generalizar resultados a la población diana o por la implicación de signos y síntomas que representen retos de diagnóstico diferencial (Argimon y Jiménez, 2013); este factor hace que el nivel de evidencia oscile entre 1b y 1c, como se muestra en la siguiente página (ver tabla 25).

Tabla 25

Clasificación de NE y GR según Oxford.

Grado de recomendación	Nivel de evidencia	Tratamiento, prevención, etiología y daño	Pronóstico e historia natural	Diagnóstico	Diagnóstico diferencial y estudios de prevalencia	Estudios económicos y análisis de decisión
A	1a	RS con homogeneidad de EC controlados con asignación aleatoria	RS de estudios de cohortes, con homogeneidad, o sea que incluya estudios con resultados comparables, en la misma dirección y validadas en diferentes poblaciones	RS de estudios diagnósticos de nivel 1 (alta calidad), con homogeneidad, o sea que incluya estudios con resultados comparables y en la misma dirección y en diferentes centros clínicos	RS con homogeneidad de estudios de cohortes prospectivas	RS con homogeneidad de estudios económicos de nivel 1
	1b	EC individual con intervalo de confianza estrecho	Estudios de cohortes individuales con un seguimiento mayor de 80% de la cohorte y validadas en una sola población	Estudios de cohortes que validen la calidad de una prueba específica, con estándar de referencia adecuado (independientes de la prueba) o a partir de algoritmos de estimación del pronóstico o de categorización del diagnóstico o probado en un centro clínico	Estudio de cohortes prospectiva con buen seguimiento	Análisis basado en costes o alternativas clínicamente sensibles; RS de la evidencia; e incluyendo análisis de la sensibilidad
	1c	Eficiencia demostrada por la práctica clínica. Considera cuando algunos pacientes mueren antes de ser evaluados	Resultados a partir de la efectividad y no de su eficacia demostrada a través de un estudio de cohortes. Series de casos todos o ninguno	Pruebas diagnósticas con especificidad tan alta que un resultado positivo confirma el diagnóstico y con sensibilidad tan alta que un resultado negativo descarta el diagnóstico	Series de casos todos o ninguno	Análisis absoluto en términos del mayor valor o peor valor

Fuente: Manterola & Zavano. (2009). Cómo “interpretar los niveles de evidencia” en los diferentes escenarios clínicos (p. 586).

En cualquier caso, la presente ofrece un tratamiento estadístico de las variables en estudio sujetas a un control estricto de posibles sesgos, un diseño metodológico eficiente y rigor científico para su ejecución.

2.3. Marco Conceptual (Definición de términos/Palabras Clave)

Los términos empleados en la lista de palabras clave que figuran en la presente (por excepción de: trastorno depresivo o trastorno mixto ansioso y depresivo) pueden ser hallados como parte de la lista del Medical Subject Headings (*MeSH*) de Medline, el *MeSH* es el *tesauro* más utilizado en el ámbito de la medicina, para la utilización correcta del vocabulario clínico, el cual mejora y facilita el acceso a la literatura biomédica (Argimon & Jiménez, 2013).

Psicología Clínica Basada en la Evidencia (PCBE): Sugiere un enfoque en el que se integra el conocimiento de la evidencia científica sobre los tratamientos más eficaces con la propia experiencia clínica del terapeuta a la hora de tomar decisiones sobre el tratamiento psicológico a aplicar (Ybarra, Orozco & Valencia, 2015).

Medicina Basada en la Evidencia (MBE): Supone la integración de la experiencia clínica individual con la mejor evidencia disponible derivada de la investigación científica sistemática (Ybarra, Orozco & Valencia, 2015).

Estudio de Precisión Diagnóstica (EPD): Es un tipo de estudio descriptivo que requiere una mención especial. Tratan de conocer la capacidad diagnóstica de una prueba, es decir, la capacidad para clasificar a una persona según la presencia o ausencia de enfermedad (Anza, López, González et al. 2015).

Sensibilidad Diagnóstica (Se): Se define como la probabilidad de que un individuo enfermo tenga un test positivo. La sensibilidad indica la proporción del total de enfermos que el test es capaz de detectar (CTO, 2011).

Especificidad Diagnóstica (Sp): Es la probabilidad de que un individuo sano tenga un test negativo. La especificidad indica la proporción de individuos sanos confirmados como tales por el resultado negativo del test (CTO, 2011).

Discriminación Diagnóstica: Capacidad de alguna prueba diagnóstica para diferenciar sujetos sanos versus enfermos (Cerde & Cifuentes, 2012).

Area Under the Curve (AUC): Medida refleja qué tan bueno es el test para discriminar pacientes con y sin la enfermedad a lo largo de todo el rango de puntos de corte posibles (Cerde & Cifuentes, 2012).

Estándar de Oro (Gold Standard): Es una prueba para determinar la presencia o ausencia de alguna enfermedad, sintomatología o entidad nosológica, la cual posee superior rendimiento diagnóstico (CTO, 2011).

Prevalencia (P): Es el número total de casos entre el total de la población. No hace distinción entre casos antiguos y nuevos y nos describe la situación en un momento determinado del tiempo. Su valor es más alto a mayor incidencia y duración de la enfermedad (CTO, 2011).

Valor Predictivo Positivo (VPP): Probabilidad de, habiendo dado positivo en la prueba, estar realmente enfermo. Se calcula como la proporción de individuos verdaderamente enfermos [VP] sobre el total de resultados positivos [VP+FP] (Hernández, Gil, Delgado et al. 2005).

Valor Predictivo Negativo (VPN): Probabilidad de, habiendo dado negativo en la prueba, estar realmente sano. Se calcula como la proporción de individuos no enfermos [VN] sobre el total de negativos [VN+FN] (Hernández, Gil, Delgado et al. 2005).

Curva ROC: Se entiende por curvas ROC a las características operativamente funcionales de una prueba diagnóstica, el cual constituye un método estadístico para determinar la exactitud diagnóstica de tests (Cerdeira & Cifuentes, 2012).

BDI: El Inventario de Depresión de Beck es un instrumento clínico, adaptado en su segunda versión a criterios diagnósticos del DSM-IV, el cual se encuentra orientado a la identificación de sintomatología depresiva y estimarla en una escala de gravedad que confiere sus intervalos a varios puntos de corte óptimos (Beck, Steer & Brown, 2009).

CAPÍTULO III

HIPÓTESIS

Al tratarse de un *Estudio de Pruebas Diagnósticas*, en la que se estimó la *Eficacia de Precisión Diagnóstica* entre dos pruebas, las hipótesis formuladas fueron de origen estadístico ya que se estableció un supuesto acerca de la superioridad entre dos medidas distribuidas en un solo parámetro poblacional (Gonzalo, 2013; Gordillo y Medina, 2012), por lo que las hipótesis se expresaron de la siguiente manera:

3.1. Hipótesis General:

H₀: El Inventario de Depresión de Beck (BDI-II) no demostrará parámetros significativos de precisión diagnóstica [$AUC \leq AUC(EAMD)$] para la discriminación clínica entre pacientes adultos ($\geq 18a$) que presenten depresión de aquellos que no lo presenten, respecto a la Escala de Depresión de Zung (EAMD) utilizada como Estándar de Oro.

H_a: El Inventario de Depresión de Beck (BDI-II) demostrará parámetros significativos de precisión diagnóstica [$AUC > AUC(EAMD)$] para la discriminación clínica entre pacientes adultos ($\geq 18a$) que presenten depresión de aquellos que no lo presenten, respecto a la Escala de Depresión de Zung (EAMD) utilizada como Estándar de Oro.

En la sección continua a la de errores de tipo 1 y 2 (en la píldora referente a la base teórica) se establecieron medidas de control estadística en relación al Intervalo de Confianza (IC) y al Margen de Error o Potencia de Estudio; se precisa por tanto que la comprobación o rechazo de las hipótesis se encontrarán determinadas por el grado de significación estadística (p), también denominada *nivel de significación alfa*, “la cual se define por la probabilidad máxima que se fija para que las diferencias en los resultados hallados sean debidos simplemente al azar [es decir H_0 cierta] que por convenio y de forma arbitraria, es equivalente a un valor de p del 5% (0,05), por lo que si $p < 0,05$ entonces las diferencias halladas serán reales y no producto del azar [cuanto menor sea el valor de p mayor será la evidencia contra la H_0], aceptándose la H_a , demostrando significancia estadística que rechace la H_0 ” (Madero, Pérez & San José, 2007 p. 65).

3.2. Hipótesis Específicas:

He₁: El inventario clínico de Depresión de Beck demuestra parámetros significativos de validez diagnóstica, respecto a la escala de Depresión de Zung utilizada como prueba de referencia.

He₂: El inventario clínico de Depresión de Beck expresa parámetros significativos de confiabilidad diagnóstica, respecto a la escala de Depresión de Zung utilizada como prueba de referencia.

He₃: El inventario clínico de Depresión de Beck exhibe parámetros significativos de reproductibilidad diagnóstica, respecto a la escala de Depresión de Zung utilizada como prueba de referencia.

3.3. Variables (Definición Conceptual y Operacional)

Síntesis operativa de instrumentos como variables.

<i>Variable</i>	<i>Definición Conceptual</i>	<i>Definición Operacional^(*)</i>	<i>Dimensión</i>	<i>Sub-Dimensiones</i>	<i>Categorías/Criterios de medición</i>	<i>Escala de Medición</i>
Variable Dependiente: Inventario de Depresión (BDI-II) de Beck	BDI-II: El Inventario de Depresión de Beck es un instrumento clínico, adaptado en su segunda versión a criterios diagnósticos del DSM-IV, el cual se encuentra orientado a la identificación de sintomatología depresiva y estimarla en una escala de gravedad que confiere sus intervalos a varios puntos de corte óptimos (Beck, Steer & Brown, 2009).	En el presente estudio la definición operacional se orientó a describir propiedades estadísticas de rendimiento: <i>El Inventario de Depresión de Beck es un instrumento con buenas propiedades psicométricas, puesto que aporta resultados válidos, confiables y reproducibles, al momento de detectar severidad sintomatológica depresiva, con base en sus puntos de corte operativo.</i>	Validez	Sensibilidad [92%] Especificidad [82%]	<u>Rangos determinados por sus mejores puntos de cortes para la discriminación diagnóstica.</u> (Beck, Steer & Brown, 2009). Depresión mínima (≤ 10). Depresión leve (11-17). Depresión moderada (18-29). Depresión severa ($30 \geq$).	Cualitativa ordinal/No paramétrico
			Confiabilidad	Alfa de Cronbach [0.89]		
			Reproductibilidad	Índice Kappa [0.45]		
			Eficacia	Área Bajo la Curva [0.83/como Gold Standard: SCID-I]		

<i>Variable</i>	<i>Definición Conceptual</i>	<i>Definición Operacional(*)</i>	<i>Dimensión</i>	<i>Sub-Dimensiones</i>	<i>Categorías/Criterios de medición</i>	<i>Escala de Medición</i>
Variable Independiente: Escala de Zung de Depresión (EAMD)	EAMD: La escala de autoevaluación de la depresión (EAMD) se encuentra destinada a la estimación numérica o cuantitativa de la depresión como condición clínica, en 4 rangos de severidad clínica.	En el presente estudio la definición operacional está destinada a describir propiedades estadísticas de rendimiento: <i>La Escala de Depresión de Zung, es un instrumento que demuestra propiedades psicométricas óptimas, a razón de aportar resultados válidos, confiables y reproducibles, al detectar la severidad sintomatológica relacionada a la depresión, con base en sus rangos diagnósticos.</i>	Validez	Sensibilidad [94.7%/IC:90.5%-99.7%] Especificidad [67.0%/IC:56.3%-76.3%]	<u>Rangos no fundamentados por puntos de corte para su grado de discriminación diagnóstica.</u> Depresión normal (≤ 40). Depresión mínima a moderada (41-47). Depresión marcada a severa (48-55). Depresión en grado máximo ($56 \geq$).	Cualitativa ordinal/No paramétrica.
			Confiabilidad	VPP [37.5%] VPN [98.4%] Alfa de Cronbach [0.80] Test Re-Test [0.84]		
			Reproductibilidad	Índice Kappa [0.39]		
			Eficacia	Área Bajo la Curva [0.85/como Gold Standard: SCID-I]		

CAPÍTULO IV

METODOLOGÍA

4.1. Método de Investigación

El presente estudio fue desarrollado en el marco del *método científico*, en tal sentido, el proceso del mismo fue guiado de manera sistematizada, organizada y objetiva, con la finalidad de describir un fenómeno en contexto relacionado al diagnóstico, habiendo estudiado los parámetros de precisión diagnóstica de un par de instrumentos frecuentemente utilizados en la práctica clínica y de la salud en Psicología, en consecuencia, determinando la eficacia de los mismos. Como se acaba de señalar, la presente se valió del método científico, pero más específicamente uno de carácter *hipotético-deductivo*, por lo que para dar explicación al fenómeno planteado, se formularon inicialmente premisas asociadas al problema (diagnóstico) para consecutivamente contrastarlos con la realidad (Sánchez y Reyes, 2009), que para el caso en cuestión, se mantuvo la premisa de que el BDI-II como inventario clínico para la Depresión demostraría parámetros de precisión diagnóstica superiores a los que mostraría la escala de Depresión de Zung, parámetros diversos condensados en un criterio unificador, a señalar, el AUC (Area Under the Curve: Área bajo la curva), por lo que tras la confrontación de la realidad a través del tratamiento estadístico pertinente, se pudo aceptar la hipótesis alterna, la cual coincide con la premisa señalada inicialmente.

Por otro lado, el presente estudio de eficacia diagnóstica fue realizado en un contexto clínico, por lo que la elección metodológica más idónea y pertinente se encontró relacionada a una propia de las ciencias de la salud, por lo que la presente se fundamentó bajo principios de la Investigación Clínica y Epidemiológica para resolver en la medida de sus alcances y limitaciones científicas la incertidumbre asociada a la actividad más frecuente en la práctica clínica, a saber, el diagnóstico y, como parte del fundamento metodológico referencial para la presente, se asumió enteramente el aporte para el diseño metodológico desarrollado por Josep Argimon y Josep Jiménez, los cuales integran la estrategia de investigación y demás en su libro *“Métodos de Investigación Clínica y Epidemiológica”*, referencia que dicho sea de paso, es frecuentemente citada en estudios de alto rigor científico (Argimon y Jiménez, 2013), la cual es entendida como un enfoque científico diseñado para que profesionales de la salud elaboren respuestas a los principales problemas que surgen en el contexto de trabajo clínico (etiología, diagnóstico, pronóstico, tratamiento, prevención, promoción, entre otros), con la finalidad de describir tales fenómenos, hallar y valorar asociaciones entre exposición y resultado o incluso contrastar premisas experimentalmente, todo ello, de la mano a la estadística epidemiológica, la cual representa la principal herramienta en Investigación Clínica con la que se controla el azar, sesgo o confusión de los resultados que puedan obtener respecto a la formulación de estos fenómenos, aportando consecuentemente evidencia al juicio clínico de los profesionales frente a la actuación con pacientes o el manejo de la práctica clínica, permitiendo tener información sobre la utilidad y eficacia de procedimientos diagnósticos, terapéuticos, preventivos, así como sobre la etiología y factores de riesgo en problemas de salud (García, 2012; Argimon & Jiménez, 2013).

Sus clasificaciones en resumidas cuentas (puesto que no es de obligación fundamentarla detalladamente) se orientan a la elaboración de diseños (*basados en el control de la asignación de los factores de estudio, es decir, Vi*) observacionales (EO) o experimentales (EE) capaces de resolver la problemática científica producida en ambientes clínicos y de la salud (los cuales comúnmente giran en torno a la incertidumbre del diagnóstico normal o anormal y su significado que representa el resultado para el paciente) (Manterola, 2009).

Teniendo el primer diseño el objetivo de la observación y el registro de los acontecimientos clínicos sin intervención alguna en su curso natural y el segundo diseño a manipular los acontecimientos clínicos y determinar el efecto de dicha intervención en su curso natural (Manterola, 2009).

Argimon y Jiménez añadieron 3 criterios o ejes más de clasificación, los cuales pueden caracterizar a un estudio por su *finalidad*, pudiendo ser *analítico* (evalúa una presunta relación causal entre un factor y un efecto o respuesta) o *descriptivo* (describe parámetros que caracterizan a poblaciones por el factor de estudio), por otro lado, de acuerdo a la *secuencia temporal*, pudiendo ser *transversal* (los estudios con datos de cada individuo que representa un momento determinado de tiempo) o *longitudinal* (existencia de periodos de tiempo entre las variables en estudio) y finalmente por su *inicio del estudio en relación con la cronología de los hechos*, pudiendo ser *prospectivos* (estudio cuyo inicio es anterior a los hechos estudiados) o *retrospectivos* (estudios cuyo diseño es posterior a los hechos estudiados) y es en función de estos principios que se estableció el tipo de estudio equivalente al de la presente (Argimon & Jiménez, 2013).

4.2. Tipo de Investigación

El tipo de investigación es coherente al *sustantivo descriptivo*, puesto que se limita a trazar parámetros puramente descriptivos del problema en estudio, buscando explicar las propiedades y características más relevantes del fenómeno analizado sujeto a una población objetivo (Argimon y Jiménez, 2013; Hernández, Fernández y Baptista, 2014; Sánchez y Reyes, 1998); en el caso de la presente, se describió parámetros de precisión diagnóstica en dos instrumentos clínicos para valorar depresión como problema de salud, parámetros que se obtuvieron producto del análisis estadístico y epidemiológico, la cual facilitó la distinción respecto a la eficacia diagnóstica entre ambas pruebas.

Como parte de la metodología clínica y epidemiológica, se denomina Estudios de Validez Diagnóstica, pero que de acuerdo a los Grados de Recomendación (GR) científica se denominan correctamente como Estudios de Precisión Diagnóstica, el cual forma parte de los estudios descriptivos u observacionales en la Investigación Clínica. Con el que se busca determinar parámetros de precisión diagnóstica y compararla respecto a un Gold Standard, sin la necesidad de hallar alguna forma de relación causal entre un factor de estudio y una variable de respuesta, aportando en última instancia mejoras operativas para su instrumentación (González & Velarde, 2012).

4.3. Nivel de Investigación

El nivel de investigación clínica que mantiene la presente correspondería al nivel *descriptivo* dentro los 6 alcances que tiene la metodología clínica y epidemiológica, ya que ésta en específico se caracteriza “desde el punto de vista cognoscitivo a describir fenómenos clínicos en una circunstancia temporal y geográfica temporal y desde el punto de vista estadístico su propósito es estimar parámetros” (Hernández, Fernández y Baptista, 2014), sin embargo, su

estadística consiste básicamente en una de tipo univarida, la cual no es exclusiva de la presente investigación, ya que se aplicó un conjunto de técnicas estadísticas como Análisis de Distribución Diagnóstica, Alfa de Cronbach, Índice Kappa, Kolmogorov-Smirnov, Curvas ROC y Signos de Wilcoxon; estadísticos propios de un nivel de investigación *predictiva* dentro la Investigación Clínica, la cual se encarga de estimar parámetros probabilísticos respecto a un evento relacionado a la salud (Donis, 2013), tal y como sucede en la presente, ya que las estimaciones de parámetros de precisión diagnóstica están directamente relacionadas con determinar la eficacia que las pruebas en estudio mantienen al calcular la probabilidad de presencia o ausencia de la condición clínica en estudio, lo que se estudia aquí no es la depresión, sino la eficacia que los instrumentos demuestran para su detección probabilística, por lo que el nivel al que corresponde es al *predictivo*.

4.4. Diseño de Investigación

La estrategia de investigación por la cual se diseñó la presente investigación engloba el conjunto de procedimientos, métodos y técnicas, mediante el cual se seleccionó a los casos, se recogió los datos, se realizó si análisis y contigua interpretación, en este sentido equivale a un diseño: *Cuantitativo-no experimental*, por ende *observacional* ya que se observó y registró acontecimientos relacionados a la detección clínica de depresión, sin intervenir de alguna forma en el curso natural del mismo, *descriptivo* puesto que se describió parámetros probabilísticos de precisión diagnóstica, de secuencia temporal *transversal* ya que tal descripción de parámetros se realizó en un momento específico del tiempo y con inicio de estudio en relación con la cronología de los hechos de tipo *prospectivo* puesto a que el inicio del presente estudio fue anterior a los hechos estudiados, habiendo sido recogidos a medida que se iba administrando los instrumentos de detección clínica (Argimon y Jiménez, 2013); por lo tanto, su esquema de diseño sería el siguiente:

M ---- O₁

4.5. Población y Muestra

La población diana (a la cual se generaliza los resultados de la misma) está constituida por todos los pacientes que acuden al servicio de Salud Mental del Hospital RDCQ Daniel Alcides Carrión ubicada en la ciudad de Huancayo, hospital de categoría III-E (de atención especializada), responsable de asumir la principal carga de enfermedad en la región Junín; de la cual se estima que atiende en promedio a 65,000 casos como población beneficiaria general anualmente, en este sentido, esta última cifra sería la que representaría a la población intrahospitalaria, por otro lado y más específicamente al servicio de Salud Mental, se estima que dicho servicio atiende en promedio a 3,000 pacientes anualmente y que casi una tercera parte de la población asistida por este servicio padece algún trastorno del estado de ánimo, considerando a este grupo clínico como el más frecuente en el servicio; las características sociodemográficas que registra el servicio describe con mayor frecuencia a pacientes adultos, con prevalencia en mujeres, de estrato socioeconómico media-baja, secundaria completa como grado de instrucción promedio, de estado civil conviviente y/o soltero(a) y respecto a la clínica, pacientes pluripatológicos (comorbilidad frecuente), con cuadros clínicos prolongados (en minoría cuadros agudos) y de casos con complicado acceso a la salud (sinónimo también de baja adherencia).

La muestra estuvo constituida por 210 pacientes adultos con edad mayor o igual a 18 que acuden al servicio de Salud Mental del Hospital RDCQ Daniel Alcides Carrión, seleccionados de manera no probabilística por casos consecutivos, el cálculo de tamaño muestral estuvo sujeto a consideraciones estadísticas y metodológicas coherentes con el tipo de investigación, dicho proceder se describe a continuación:

El cálculo del tamaño muestral para este tipo de estudios aún no se ha estandarizado respecto a su procedimiento más adecuado y, aunque la literatura señale que no es indispensable un cálculo muestral de extrema precisión para este tipo de estudios, los manuales de Investigación Clínica y los estándares internacionales de calidad para el estudio de pruebas diagnósticas (ya sea STARD o QUADAS) sugieren una determinación muestral en función de la prevalencia estimada de la enfermedad [especialmente si el estudio se diseña para seleccionar una sola muestra para el estudio; siendo este el caso de la presente investigación] en la población, el valor esperado de sensibilidad y/o especificidad, el IC y la Precisión Absoluta (PA) aceptada por el estudio (Gonzalo, 2013; González & Velarde, 2012).

Sin embargo, Argimon y su equipo de investigadores señalan que al tratarse de un estudio de pruebas diagnósticas en la que se calculan fundamentalmente parámetros del diagnóstico que reflejan proporciones, como lo son la sensibilidad y especificidad, el cálculo del tamaño muestral adecuado también correspondería con la estimación de una proporción (Argimon y Jiménez, 2013), en la que hay que definir sólo el IC, la PA, el tamaño poblacional y la proporción esperada al 50%, sin tener en cuenta la sensibilidad y especificidad esperada como en el procedimiento señalado con antelación. Por lo que en la presente investigación se estimó el cálculo del tamaño muestral como estipula el primer procedimiento recomendado, bajo los siguientes criterios estadísticos y metodológicos:

Consideraciones estadísticas:

- Primero: Intervalo de Confianza (IC) al 95% por recomendación de la lista de comprobación STARD para estudios de validez diagnóstica (intervalo para el control de error alfa recomendado) (González & Velarde, 2012).
- Segundo: Margen de error del 10% y PA del 90% (valor de PA sugerido para el control de error beta) (Gonzalo, 2013).
- Tercero: Los parámetros de Se (sensibilidad) y Sp (especificidad) esperados basados en estudios de validez diagnóstica del método de referencia previo que demostró mejor precisión [Depresión= Se: 92,0%; Sp: 71,0%] (Campo et al. 2005; Sanz, 2014).
- Cuarto: La proporción esperada del 50% (por desconocerla con certeza).
- Quinto: El cálculo de la prevalencia se realiza gracias a los reportes de morbilidad a los que se pudo tener acceso por parte de la Oficina de Estadística del Hospital RDCQ “Daniel Alcides Carrión, de los que se estima la prevalencia en el servicio de psicología y psiquiatría equivalente en: Depresión= 22,5% (Cadenas H., 2018).

De esta manera el cálculo del tamaño muestral para el presente estudio de precisión diagnóstica se determinó gracias al Programa Para el Análisis Epidemiológico de Datos Tabulados EPIDAT; Versión 3.1; cabe precisar que del cálculo muestral basado en la prevalencia de la enfermedad se establecen parámetros de precisión del estudio (tamaño muestral para posibles márgenes de error):

Precisión (%)	Total
5,000	507
10,000	129

Figura 13 Cálculo del tamaño muestral basado en la Prevalencia (Depresión).

Fuente: Cadenas, H. (2018). Recurso Personal

De la figura relacionada al cálculo del tamaño muestral para el conjunto clínico denominado depresión, se estimó el parámetro muestral con un IC del 95%, los parámetros discriminativos primarios como lo son la Sensibilidad y Especificidad ($Se=92\%$ y $Sp=71\%$) en este caso no esperada sino determinada por estudios previos lo cual aporta mayor consistencia para su cálculo, dejando de lado la expectativa inferencial del investigador, también se consideró la prevalencia de la enfermedad hallada en el servicio de Salud Mental del Hospital RDCQ Daniel Alcides Carrión para el 2017 y 2018, sin embargo, y es en este punto en el que se brindará mayor aclaración, puesto que al hablar de precisión diagnóstica hablamos del margen de error con el cual expresaremos los resultados de la presente investigación, siendo la precisión absoluta el porcentaje complementario al margen de error, es decir, si planteamos un margen de error del 5% [Precisión Absoluta del 95%], queremos en realidad expresar que, por ejemplo, de 100 pruebas diagnósticas en las que se analizó parámetros reales (no producto del azar), el estudio la analizaría como tal en un promedio de 95 veces.

Entonces debido a la significancia de la prevalencia de la enfermedad para calcular el tamaño muestral, la presente otorgó un nivel de precisión de estudio del 10% ya que como vemos en la tabla 26 mientras el margen de error del estudio va incrementando su magnitud, el tamaño muestral va decreciendo, a causa principal de la validez externa, es decir, la posibilidad de generalizar resultados a la población diana, sin embargo, para este tipo de estudios como ya se mencionó, dependientes de la prevalencia de la enfermedad y *un nivel de precisión del 10% e incluso del 20% (como nivel máximo de beta tolerable) se consideran aceptables*, pues describimos parámetros de precisión diagnóstica de los instrumentos y no de los casos (Gonzalo, 2013; Gómez y Sánchez, 1998), en este sentido el tamaño muestral correspondiente a este conjunto clínico equivale a 129 casos, pero se determinó 156 casos para acercarse al tamaño muestral equivalente al de un estudio con precisión del 9%, lo cual además es viable para la presente a diferencia del tamaño muestral con precisión del 5%, puesto que en la presente se determinaron parámetros de precisión diagnóstica para dos instrumentos clínicos, por tanto, ameritó mayor moderación con el método para calcular el tamaño muestral.

Tabla 26
Análisis de precisión y equivalencia muestral (Depresión).

Precisión (%)	Total
5,000	507
6,000	352
7,000	258
8,000	201
9,000	156
10,000	129

Fuente: Cadenas, H. (2018). Recurso Personal (Cálculo EPIDAT).

Teniendo en cuenta por lo tanto como *población diana* (Población a la que se desea generalizar los resultados): pacientes sujetos a condiciones clínicas asociadas a depresión y

como *población de estudio* (Población definida por los criterios de selección) a los individuos definidos por las consideraciones metodológicas.

Control estadístico de las no respuestas (procedimiento que fue realizado en el transcurso de la ejecución de la presente investigación por considerarse necesario) con la que se evitó el sesgo de estimación de resultados, a través de la cual se procuró conseguir el mayor % de respuestas posibles, considerando como criterio de exclusión a las no respuestas, además de haber corregido la proporción de las no respuestas, las pérdidas o los abandonos (de prueba a prueba), ampliándose el número de sujetos calculados a través de la siguiente fórmula (Argimon & Jiménez, 2013):

$$N_a = N [1/1 - R] = \text{Lo que llevó a considerar 210 casos, reduciendo el valor de PA}=8\%.$$

Donde: N fue el número de sujetos teórico, N_a el número de sujetos ajustados y R la proporción esperada de pérdida equivalente al 5%. Asegurando de este modo la potencia estadística deseada por la presente, ajustando un tamaño muestral final de 210 pacientes.

Consideraciones metodológicas:

Respecto a la determinación de la población, los individuos que se estudiaron no debían diferir de aquellos individuos que queremos aplicar la prueba nueva en la práctica clínica (Rodríguez, 2013).

Para evitar el sesgo de espectro se incluyó no sólo a los individuos específicamente no afectados por la enfermedad, sino a pacientes que representen retos o problemas de diagnóstico diferencial (Argimon & Jiménez, 2013; Rodríguez, 2013).

Se incluyeron también pacientes con diferentes grados de severidad de la enfermedad y no sólo pacientes con cuadros severos o subclínicos, para evitar sesgos de sobreestimación y subestimación de la capacidad discriminativa de la nueva prueba (Rodríguez, 2013; González & Velarde, 2012).

Tipo de Muestreo:

La estrategia de selección se orientó a estudiar sólo una muestra debido a sus ventajas prácticas (lo cual vuelve al estudio descriptivo a diferencia de las selección de dos muestras, compuestos por casos y controles volviendo al estudio analítico), las cuales se basan esencialmente en que se dispone de un amplio espectro de la enfermedad, además de incluir a pacientes que representen retos para su diagnóstico, ya sea por atipicidad sintomatológica, comorbilidad, dualidad, etc. y además que reduce considerablemente las desventajas que presenta esta estrategia ya que los instrumentos en estudio no se consideran costosos o invasivos (Argimon & Jiménez, 2013).

Respecto al tipo de muestreo se consideró proceder con un *muestreo no probabilístico* (muestra apareada), *con selección de casos consecutivos*, para no diferir sustancialmente de la población diana a la que se quiere generalizar los parámetros de desempeño diagnóstico que demuestren las pruebas, además de coincidir metodológicamente con las recomendaciones realizadas por las jerarquías de evidencia y grados de recomendación (comparación independiente y enmascarada de un espectro de pacientes consecutivos) (Argimon & Jiménez, 2013; Manterola, Asenjo & Otzen, 2014; Díaz, Reyes, López et al. 2006).

La *población diana* se contextualizó a la población intra-hospitalaria del Hospital RDCQ Daniel Alcides Carrión y con ello a toda la población extra-hospitalaria homóloga a las mismas características sociodemográficas que caracteriza al tipo de paciente típico que acude a consulta en este Hospital. La *población de estudio* estuvo contextualizada por pacientes que acudieron al Servicio de Salud Mental de dicho Hospital, conformado por 210 casos en total, delimitados por los criterios de selección que se describen a continuación.

Criterios de inclusión y exclusión:

Los criterios de selección no fueron estrictos para este tipo de estudios, puesto que la actividad diagnóstica para conjuntos clínicos no está perfectamente delimitada en su expresión sintomatológica, en este sentido las manifestaciones sintomatológicas asociadas a la depresión se hallaron como espectro en diversos trastornos del comportamiento y, en lo que se trató la presente fue estudiar una muestra que no difiera de la población consultante en la práctica clínica habitual, por lo que este fue el *criterio de inclusión* principal. *De exclusión*: Pacientes sujetos a tratamiento farmacológico, pacientes evaluados con anticipación o pacientes con cuadros clínicos lejanamente relacionados a los conjuntos clínicos en estudio.

4.6. Técnicas e Instrumentos de Recolección de Datos

Patrón de oro:

Se empleó como prueba de referencia o gold standard la Escala de Autoevaluación de Zung para depresión (EAMD), en razón a su desempeño diagnóstico demostrado por diversos estudios precisados a lo largo de la presente investigación, además de ser la escala utilizada como primera línea alternativa de evaluación clínica en el contexto de investigación (sugerida en protocolos por el Minsa para todo el Perú), volviéndola como la más importante prueba después de la Escala de Depresión de Hamilton (considerada como gold standard en otros

países como Argentina) (Montenegro y Kilstein, 2014) como parte de la “*jerarquía de pruebas complementarias*” (Farreras y Rozman, 2012, p.5) utilizadas en salud mental.

Prueba en estudio:

La elección del Inventario de Depresión de Beck en su segunda versión como nueva prueba en estudio de su eficacia, estuvo fundamentada por ser un instrumento sujeto a criterios diagnósticos DSM, además de ser un instrumento actualizado, con un marco teórico de la cual se dispone sustento científico (modelo cognitivo) e incluso de raíces conductuales que aportan argumentos experimentales al momento de explicar, evaluar o tratar la depresión.

4.7. Técnicas de Procesamiento y Análisis de Datos

El procesamiento de datos se realizó en dos momentos específicos del estudio, el primero después de la ejecución de la *prueba piloto* y la segunda después de *la ejecución del estudio propiamente dicho*.

Con la prueba preliminar o piloto, se evaluó dos dimensiones de la confiabilidad para la nueva prueba antes de haber evaluado su desempeño diagnóstico, es evaluó en ese sentido, la consistencia interna de sus resultados (escala completa) y la concordancia inter-examinadores (teniendo en cuenta como segundo examinador al juicio clínico de los especialistas), los cuales estuvieron sujetos a procedimientos estadísticos de tratamiento correspondientes con Coeficiente de Alfa de Cronbach (Método: Varianza de ítems) y el Índice de Kappa (Método: Acuerdo entre examinadores) respectivamente, sus resultados se describen en la sección correspondiente.

La ejecución del estudio propiamente dicho, implicó en primera instancia la descripción estadística de características sociodemográficas de la muestra, alternadamente se administró el Gold Standard a cada paciente de los 210 casos, seguido de la administración de la nueva prueba, con lo que se determinó posteriormente las distribuciones diagnósticas, las cuales fueron la base para todo el análisis ulterior, delimitando de tal modo los parámetros de validez, confiabilidad y reproductibilidad diagnóstica, así como el análisis de Curvas ROC, la comparación de eficacia por el AUC que cada prueba presentó.

Toda la actividad relacionada a emplear técnicas de procesamiento y análisis de datos se realizó de la siguiente manera: En principio se diseñó la base de datos en el software para procesamiento de datos estadísticos SPSS en su versión más actualizada (24.0.0; Desarrollado por la Universidad de Chicago), en el que se tabuló los resultados relacionados al análisis de Consistencia Interna (Alfa de Cronbach) y Concordancia Inter-Examinadores (Índice de Kappa) para los 50 casos que conformaron la prueba piloto. Para la estadística descriptiva en la que se describió características sociodemográficas, se utilizó el SPSS con el que se determinaron medidas de tendencia central para la variable edad, respecto a la variable grado de instrucción, sexo y estado civil sólo se determinó el valor más frecuente en la muestra (moda), para al análisis de frecuencias de la muestra se empleó el software Excel en su versión del 2016. El cálculo de parámetros de distribución diagnóstica fue desarrollado en el SPSS, los parámetros de validez, confiabilidad y reproductibilidad fueron calculados en el programa para el análisis epidemiológico de datos tabulados EPIDAT (desarrollado por la OPS), siendo utilizado en sus dos versiones (3.1 y 4.1) puesto que con cada uno se accedió a un tipo diferente de análisis estadístico. Para el análisis de Curvas ROC se formuló inicialmente utilizar EPIDAT, sin embargo, en el transcurso del estudio se pudo constatar que no presentaba

disponible el análisis de Curvas ROC no paramétricas, por lo que se optó utilizar el SPSS para este análisis y finalmente también para el test estadístico de Rangos de Wilcoxon.



Fuente: Cadenas, H. (2018). Recurso Personal (EPIDAT 3.1).

Figura 14 Software para el análisis epidemiológico EPIDAT.

Comprobación de Hipótesis

El tratamiento estadístico para la comprobación de la hipótesis en el presente estudio se partió inicialmente por el análisis de Curvas ROC para pruebas no-paramétricas y la estimación del intervalo de confianza para el análisis del Área bajo la curva (AUC) que permitirá rechazar o aceptar la hipótesis, en función a la eficacia de precisión del nuevo procedimiento diagnóstico (nueva prueba en estudio), ya que posteriormente se realizó el análisis comparativo respecto a la eficacia diagnóstica del BDI-II y el EAMD, utilizando el AUC como índice general de medida de la eficacia (sujeto a sus intervalos de confianza como control del error aleatorio y como criterio para la identificación de la diferencia más pequeña que ofreció el valor clínicamente relevante de los resultados).

Además de todo ello, se empleó la prueba no paramétrica de los Rangos de Wilcoxon para escalas ordinales, con la que se permitió afianzar el contraste de hipótesis, en el sentido de que aportó la significancia estadística (valor de $p < 0,05$ fijado a priori para evitar el sesgo de conveniencia) suficiente para determinar si realmente existieron diferencias entre la eficacia

entre ambos instrumentos (en su precisión diagnóstica producto del análisis de sus medidas o resultados) o si sólo fue producto del azar, controlando de ese modo el error tipo I (concluir que existen diferencias entre dos instrumentos cuando en realidad no existen) y el riesgo alfa (probabilidad de cometer el error de tipo I) (Burgueño, García & González, 1995; Argimon & Jiménez, 2013).

4.8. Aspectos Éticos de la Investigación

Se mantuvo la confidencialidad de resultados, salvaguardando el bienestar e integridad de los pacientes, tomando en cuenta su participación de manera anónima, en el transcurso del estudio algunos pacientes accedieron a brindar sus datos personales, algunos se limitaron a brindar su número de DNI y un grupo considerable se abstuvo incluso a firmar el consentimiento informado que certificara su participación voluntaria en el estudio, a pesar de ello, participaron en el estudio sin estar sujetos a ningún documento de confidencialidad, accediendo a que sus resultados fueran publicados, puesto que al no firmar los consentimientos informados, no habría problema en que se utilice sus resultados para lo que se considere necesario, siempre y cuando termine aportando mejoras para el servicio que como usuarios reciben.

Los documentos relacionados al consentimiento informado para la participación voluntaria y la publicación de resultados se muestran en el Anexo 7. Los cuales fueron diseñados bajo criterios de investigación clínica y con referencia a un estudio de eficacia diagnóstica (Montenegro & Kilstein, 2014).

CAPÍTULO V

RESULTADOS

5.1. Descripción de Resultados

En el siguiente apartado se exponen detalladamente los resultados adquiridos a lo largo del presente estudio, resultados que además de describir el proceder científico, fundamentarán la resolución del problema de investigación científica que se planteó en la presente investigación, los cuales son presentados a través de la elección de procedimientos metodológicamente pertinentes y sus consecuentes resultados estadísticamente significativos.

Resultados Preliminares: *Prueba Piloto*

Como parte del criterio preliminar al realizar un *Estudio de Precisión Diagnóstica* y el rigor metodológico del mismo, sugiere no iniciar el análisis sin antes evaluar al menos la reproductibilidad que exhibe la nueva prueba en estudio (Argimon y Jiménez, 2013, p. 91), por lo que para la presente se consideró evaluar 2 dimensiones de la confiabilidad como análisis preliminar (incluido por tanto en la fase de prueba piloto) al estudio propiamente dicho.

Para ello se determinó evaluar tanto la dimensión de Consistencia Interna (método: varianza de ítems) utilizando al Alfa de Cronbach como el tratamiento estadístico y por otro lado la

dimensión de Reproducibilidad o Confiabilidad Entre-Examinadores (método: acuerdo entre-examinadores) optando por el Índice de Kappa como procedimiento estadístico. En tal sentido, se realizó el análisis preliminar sobre la base de 50 casos y los hallazgos se exponen a continuación:

Alfa de Cronbach: Coeficiente de Consistencia Interna

En la tabla 27 se describe el coeficiente de consistencia interna del resultado que exhibe el Inventario de Depresión de Beck, dicho coeficiente denota la consistencia interna que para la presente fue de la escala completa, no se presentan resultados por dimensiones o por ítems de la prueba puesto que el procedimiento estadístico fue escogido sólo para la prueba piloto y no para un estudio puramente relacionado a la confiabilidad.

Tabla 27
Coeficiente de Consistencia Interna (BDI-II).

Alfa de Cronbach			
Pruebas de Estudio	Herramientas de Detección Clínica	Coeficiente de Fiabilidad	Grado de Homogeneidad/Criterios de George y Mallery (2003)
	Inventario de Depresión de Beck (BDI-II)	,793	Aceptable.

Fuente: Cadenas, H. (2019). Recurso Personal (Cálculo SPSS).

En tal sentido, se observó que el Coeficiente de Consistencia Interna fue equivalente a 0,79 lo cual por convenio al juicio de expertos (George y Mallery, 2003) sugiere aprobación estadística, concluyendo por tanto que el Inventario de Depresión de Beck se muestra como una prueba Aceptable en lo que a su Consistencia Interna refiere.

Índice de Kappa: Coeficiente de Concordancia Inter-Examinadores

En la tabla 28 se describe el coeficiente de concordancia inter-examinadores del resultado que exhibe el Inventario de Depresión de Beck, para el cual se asumió como patrón de referencia al juicio clínico de los especialistas, con la intención de estimar el grado de replicación que presentaría el BDI-II.

Tabla 28
Coeficiente de Concordancia Inter-Examinadores (BDI-II).

Índice Kappa			
Pruebas de Estudio	Herramientas de Detección Clínica	Coeficiente de Concordancia Inter-Examinadores	Fuerza de Concordancia/Criterios de Tornimbeni, Pérez y Olaz (2008)
	Inventario de Depresión de Beck (BDI-II)	,712	Buena.

Fuente: Cadenas, H. (2019). Recurso Personal (Cálculo SPSS).

De manera que, se observó que el Coeficiente de Concordancia Inter-Examinadores fue equivalente a 0,71 lo cual por convenio al juicio de expertos (Tornimbeni, Pérez y Olaz, 2008) sugiere aprobación estadística, concluyendo por tanto que los resultados del Inventario de Depresión de Beck replican la consideración clínica de los especialistas en un Buen grado.

Por lo demostrado, se consideró oportuno iniciar con el estudio propiamente dicho, ya que el Inventario de Depresión de Beck demostró propiedades estadísticas suficientes en relación a su Consistencia Interna y Concordancia Inter-Examinadores.

Análisis Sociodemográfico: Estudio de Población

Como parte de la propuesta de investigación, se planteó realizar el análisis sociodemográfico, con la finalidad de describir parámetros estadísticos que caracterizan a la población en estudio y en ese sentido tener a consideración la población diana a la cual se

generalizarán los resultados del presente estudio. Se describen a continuación las principales características estadísticas de la muestra estudiada por la presente:

Tabla 29

Estadística Descriptiva Edad (+DE).

	N	Mínimo	Máximo	Media	Desviación estándar
	Estadístico	Estadístico	Estadístico	Estadístico	Error estándar
Edad/Casos	210	18	74	36,78	1,049
					Estadístico
					15,196

Fuente: Cadenas, H. (2019). Recurso Personal (Cálculo SPSS).

Tabla 30

Estadística Descriptiva Media-Mediana-Moda.

	Edad/Casos	GI	Sexo/Casos	EstadoCivil
N	Válido	210	210	210
	Perdidos	0	0	0
Media	36,78	Moda (Valor)		
Mediana	34,00	4 (Sup)	2 (Mujer)	1 (Soltero)
Moda	21			

Fuente: Cadenas, H. (2019). Recurso Personal (Cálculo SPSS).

La muestra estuvo constituida por 210 casos no probabilísticos, adultos (18>) elegidos por casos consecutivos que acudieron al servicio de Salud Mental del Hospital Regional D.C.Q. Daniel Alcides Carrión, de los que la edad promedio de los casos correspondió a 36,78 años [DE=15,19/Error estándar=1,04], además que el 50% de casos tiene menos de 34 años y el 50% es mayor a dicha edad, finalmente, la edad más frecuente registrada en el presente estudio correspondió a 21 años, respecto al grado de instrucción (GI) el nivel Superior no diferenciado estadísticamente entre Superior completo e incompleto (C/I), respecto al sexo más frecuente por casos se identificó que corresponde al grupo de mujeres y la categoría de soltero(a) como estado civil con mayor presentación.

Tabla 31
Frecuencias de Distribución (Estado Civil/GI/Sexo).

Estado Civil			Grado de Instrucción			Sexo					
	Frx.	%		Frx.	%		Frx.	%			
Válido	Soltero(a)	82	39.0	Válido	Primaria (C/I)	24	11.4	Válido	Varón	61	29.0
	Con pareja	28	13.3		Secundaria (C/I)	60	28.6		Mujer	149	71.0
	Conviviente	47	22.4		Superior Técnico (C/I)	29	13.8		Total	210	100.0
	Casado(a)	38	18.1		Superior (C/I)	93	44.3				
	Divorciado(a)	8	3.8		Iltrado(a)	4	1.9				
	Viudo(a)	7	3.3		Total	210	100.0				
Total	210	100.0									

Fuente: Cadenas, H. (2019). Recurso Personal (Cálculo Excel).

Por otro lado, en la tabla 31 se describió la distribución estadística de categorías por variable sociodemográfica, los resultados se presentan en las siguientes figuras que contienen la distribución por recuentos y porcentajes, presentados en gráficos.

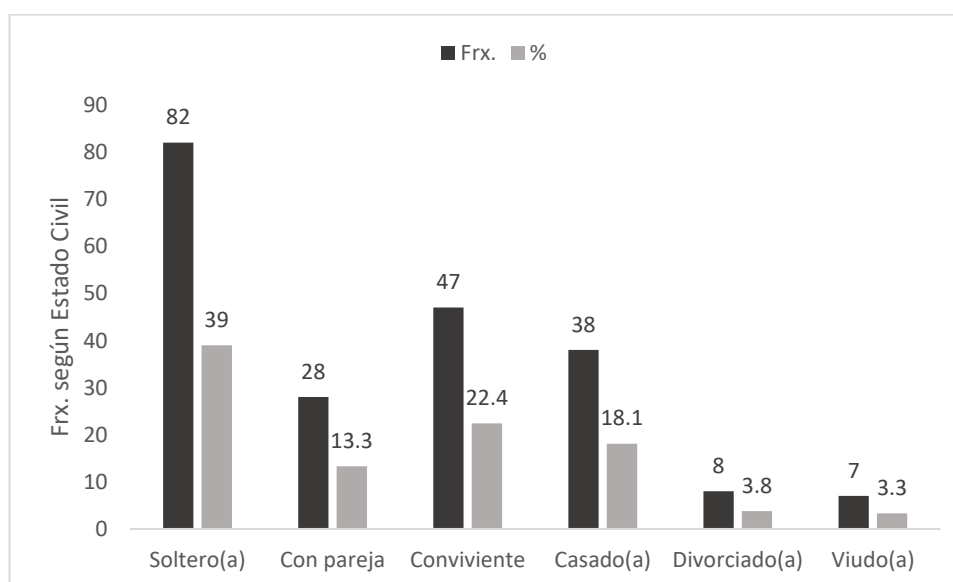


Figura 15 Frecuencia Según Estado Civil.

Fuente: Cadenas, H. (2019). Recurso Personal (Plantilla Excel).

La distribución para Estado Civil denotó un 39% de casos que se encontraron Solteros(as), siendo de tal modo el grupo estudiado con mayor frecuencia [82 casos], seguidos del grupo Conviviente, el cual representó un 22,4% [47 casos], en comparación a los dos grupos más pequeños, como los referidos a Divorciado(a) [3,8%/ 8 casos] y Viudo(a) [3,3%/ 7 casos].

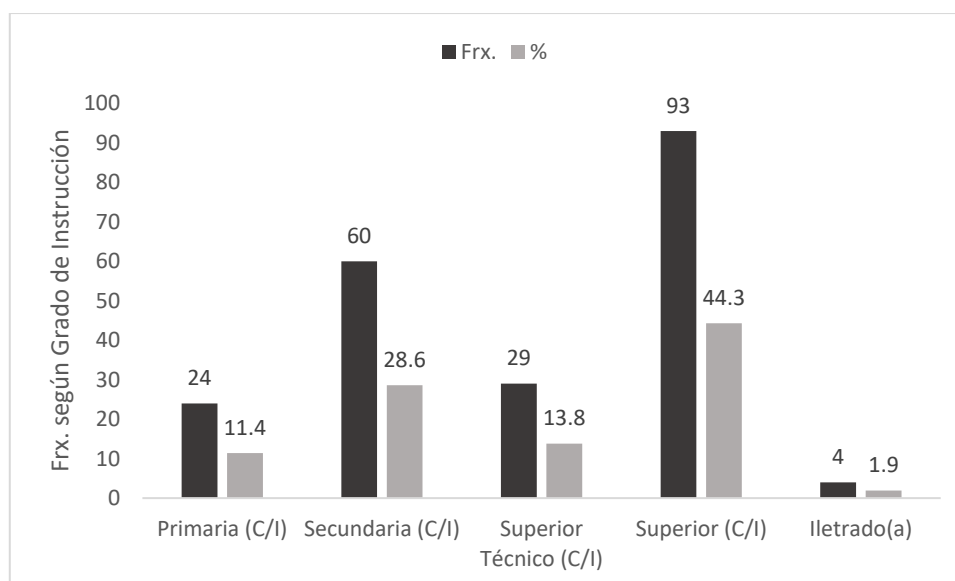


Figura 16 Frecuencia Según Grado de Instrucción.

Fuente: Cadenas, H. (2019). Recurso Personal (Plantilla Excel).

La distribución para Grado de Instrucción describió un 44,3% de casos con instrucción Superior (C/I), siendo el grado de instrucción con mayor presentación de casos, equivalente a 93, seguidamente el grupo con grado de instrucción correspondiente a Secundaria (C/I) presenta 28,2% de la frecuencia equivalentes a 60 casos del total de estudiados, siendo el segundo grupo más frecuente, en comparación al grupo de Iltrado(a) el cual fue el grupo con solo 4 casos de presentación [1,9%].

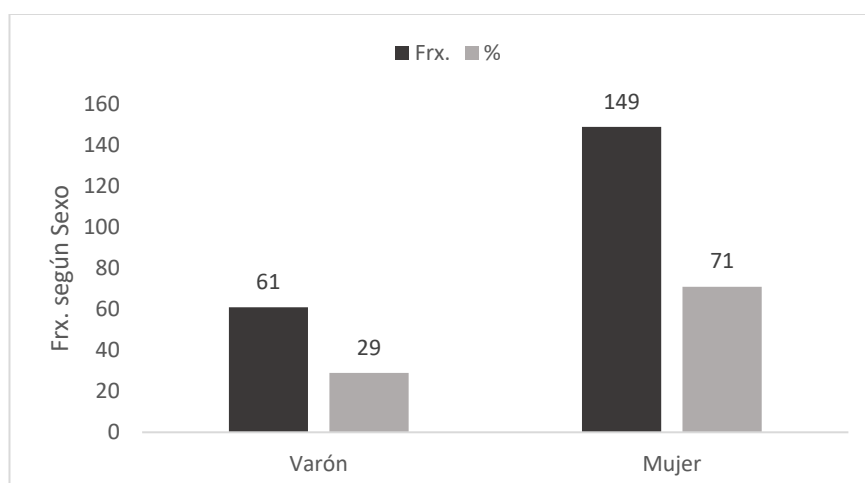


Figura 17 Frecuencia Según Sexo.

Fuente: Cadenas, H. (2019). Recurso Personal (Plantilla Excel).

Finalmente, la distribución para Sexo demostró que un gran porcentaje de la muestra estudiada corresponde al grupo de Mujeres, equivalente al 71% [149], en contraste al grupo de

Varones de los que sólo se presentó 61 casos, siendo su equivalente del 29%, pudiendo confirmar los datos epidemiológicos referidos por algunos manuales psiquiátricos o psicopatológicos que indican que la prevalencia de depresión es mayor en mujeres que varones a diferencia de su vulnerabilidad para el riesgo suicida, en la que se asocia una mayor prevalencia para varones que mujeres (Chinchilla, Correas et al. 2011).

Resolución de Problemas Específicos

En el siguiente apartado se dio resolución a los 3 problemas específicos que fueron planteados por la presente, los cuales sirvieron como fundamento preliminar para la ulterior comprobación de hipótesis, por lo que este apartado está separado por cada problema específico y su equivalente objetivo específico que argumentó el procedimiento resolutivo con la intención de comprobar las hipótesis específicas que se plantearon en la presente.

¿Describe el inventario clínico de Depresión de Beck parámetros significativos de validez diagnóstica, respecto a la escala de Depresión de Zung utilizada como prueba de referencia?

Para dar resolución a este problema específico, se formuló el siguiente objetivo específico: *Estimar* parámetros estadísticos de validez diagnóstica que aporta el Inventario de Beck para la detección clínica de Depresión, respecto a la prueba de referencia. En ese sentido, el procedimiento que se siguió fue hallar inicialmente los parámetros de distribución diagnóstica existentes entre el BDI-II y el Gold Standard presentados en un cuadro de distribución 2x2, para posteriormente determinar parámetros de validez diagnóstica, los cuales se presentan a continuación.

Tabla 32
Parámetros de Distribución Diagnóstica (EAMD/BDI-II).

IC:95.0%		EAMD		Total
		Verdadero (≥ 41)	Falso (≤ 40)	
BDI-II	Positivo (≥ 11)	73 (VP) 72,3%	28 (FP) 27,7%	101 100,0%
	Negativo (≤ 10)	12 (FN) 11,0%	97 (VN) 89,0%	109 100,0%
Total		85 40,5%	125 59,5%	210 100,0%

Fuente: Cadenas, H. (2019). Recurso Personal (Cálculo SPSS).

De la tabla 32 se pudo hallar la distribución diagnóstica que presenta el BDI-II respecto al EAMD como Gold Standard, en ese sentido se identifican 73 Verdaderos Positivos y 97 Verdaderos Negativos, siendo estos los casos válidos para el diagnóstico, en contraste a los 12 casos Falsos Negativos y 28 Falsos Positivos. A partir de estos parámetros es que se pudo hallar los siguientes parámetros: Sensibilidad (Se), Especificidad (Sp), Índice de Validez (IV), Prevalencia (P) e Índice de Youden, todos estos parámetros sujetos al Intervalo de Confianza de 95.0%. Los resultados se muestran en la siguiente figura.

	Valor	IC (95%)	
Sensibilidad (%)	85.88	77.89	93.87
Especificidad (%)	77.60	69.89	85.31
Índice de validez (%)	80.95	75.40	86.50
Índice de Youden	0.63	0.53	0.74
Prevalencia (%)	40.48	33.60	47.35

Figura 18 Parámetros de Validez Diagnóstica (EAMD/BDI-II).

Fuente: Cadenas, H. (2019). Recurso Personal (Cálculo EPIDAT 3.1).

Para dar resolución a la pregunta, se hizo necesario adquirir los parámetros de validez que exhibía el Gold Standard y de ese modo comparar su eficacia respecto a la validez diagnóstica de ambas pruebas. Es por ello que en el caso del EAMD se utilizó como Gold Standard al juicio clínico de los profesionales, debido a que al considerar al AEMD como Gold Standard

hablamos de que cuenta con parámetros discriminativos altamente significativos los cuales pueden equipararse parcialmente a la elaboración clínica de los especialistas. Sus resultados se muestran en la siguiente tabla.

Tabla 33
Parámetros de Distribución Diagnóstica (EA/EAMD).

IC:95.0%		Estado Actual del Paciente (Dx. Profesional)		Total
		Verdadero	Falso	
EAMD	Positivo (≥41)	56 (VP) 65,9%	29 (FP) 34,1%	85 100,0%
	Negativo (≤40)	25 (FN) 20,0%	100 (VN) 80,0%	125 100,0%
Total		81 38,6%	129 61,4%	210 100,0%

Fuente: Cadenas, H. (2019). Recurso Personal (Cálculo SPSS).

De esta segunda tabla se pudo inferir por mera observación que las proporciones de casos con diagnósticos inválidos es más alta considerando al juicio clínico como Gold Standard y al EAMD como simulación de nueva prueba en estudio, en este sentido se hallaran también sus parámetros de validez, los cuales se presentan a continuación.

	Valor	IC (95%)	
Sensibilidad (%)	69.14	58.46	79.81
Especificidad (%)	77.52	69.93	85.11
Índice de validez (%)	74.29	68.14	80.44
Índice de Youden	0.47	0.34	0.59
Prevalencia (%)	38.57	31.75	45.39

Figura 19 Parámetros de Validez Diagnóstica (EA/EAMD).

Fuente: Cadenas, H. (2019). Recurso Personal (Cálculo EPIDAT 3.1).

Al comparar los parámetros de validez diagnóstica que exhibieron ambas pruebas se evidenció superioridad en capacidad discriminativa para el BDI-II respecto al Gold Standard, mientras que el Gold Standard (desde ahora denominado GS) demuestra $Se=69.14\%$ y $Sp=77.52\%$ siendo la prevalencia determinada del 38.57% , el BDI-II demuestra superioridad con $Se=85.88\%$ y $Sp=77.60\%$, por lo que de entrada demuestra mejor capacidad para

discriminar pacientes afectados por la condición clínica con resultado positivo en la prueba, de tal modo incrementando su Prevalencia a 40.48%, difiriendo en 1.91%, tratándose claro está de una muestra de 210 casos, lo cual se amplificaría significativamente para muestras grandes propias de estudios epidemiológicos. En tal sentido y debido a los 2 últimos parámetros [Índice de Validez (80.95% vs 74.29%)/Índice de Youden (0.63 vs 0.47)] en los que también se encuentran diferencias significativas que denotan la superioridad del BDI-II, se da respuesta al problema específico considerando que, en efecto, el BDI-II demuestra parámetros significativos de validez diagnóstica respecto al GS (Escala de Depresión de Zung). El análisis e interpretación de lo descrito en la presente sección (resolución de problemas específicos) se precisará en la sección destinada a la Discusión de Resultados.

¿Describe el inventario clínico de Depresión de Beck parámetros significativos de confiabilidad diagnóstica, respecto a la escala de Depresión de Zung utilizada como prueba de referencia?

El objetivo específico que se planteó para dar resolución a este problema específico fue: *Estimar* parámetros estadísticos de confiabilidad diagnóstica que aporta el Inventario de Beck para la detección clínica de Depresión, respecto a la prueba de referencia. Por lo que, con base a los parámetros de validez diagnóstica previamente determinados, se halló parámetros de confiabilidad diagnóstica [Valor Predictivo +/- (VPP/VPN) y Razones de Verosimilitud +/- (LR+/LR-): ambos de carácter predictivo], los cuales se presentan como sigue.

	Valor	IC (95%)	
Valor predictivo + (%)	72.28	63.05	81.50
Valor predictivo - (%)	88.99	82.66	95.33
Razón de verosimilitud +	3.83	2.74	5.37
Razón de verosimilitud -	0.18	0.11	0.31

Figura 20 Parámetros de Confiabilidad Diagnóstica (EAMD/BDI-II).

Fuente: Cadenas, H. (2019). Recurso Personal (Cálculo EPIDAT 3.1).

En la figura 20 se describen los parámetros de confiabilidad diagnóstica que demostró el BDI-II, consideramos valores significativos tanto para los Valores Predictivos Positivos y Negativos [(+)72.28%/(-)88.99%] como para las Razones de Verosimilitud [(+)3.83/(-)0.18].

	Valor	IC (95%)	
Valor predictivo + (%)	65.88	55.22	76.55
Valor predictivo - (%)	80.00	72.59	87.41
Razón de verosimilitud +	3.08	2.16	4.37
Razón de verosimilitud -	0.40	0.28	0.56

Figura 21 Parámetros de Confiabilidad Diagnóstica (EA/EAMD).

Fuente: Cadenas, H. (2019). Recurso Personal (Cálculo EPIDAT 3.1).

Para realizar la comparación respecto al rendimiento de confiabilidad diagnóstica entre ambas pruebas, se determinó independientemente parámetros de confiabilidad diagnóstica del EAMD utilizado en simulación de nueva prueba y como GS al juicio clínico de especialistas, en tal sentido la EAMD demostró inferioridad diagnóstica en relación al Valor Predictivo Positivo [(BDI-II) 72.28% vs 65.88% (EAMD)], demostró diferencia mínima respecto al VPN (Diferencia entre BDI-II/EAMD=8.99%), respecto a las Razones de Verosimilitud se describió cambios significativos, en tal sentido y del mismo modo que para los parámetros de validez, se responde al segundo problema específico considerando que, en efecto, el BDI-II demuestra parámetros significativos de confiabilidad diagnóstica respecto al GS (Escala de Depresión de Zung=EAMD).

¿Describe el inventario clínico de Depresión de Beck parámetros significativos de reproductibilidad diagnóstica, respecto a la escala de Depresión de Zung utilizada como prueba de referencia?

Finalmente, se estableció el siguiente objetivo específico para dar resolución al problema específico señalado: *Estimar* parámetros estadísticos de reproductibilidad diagnóstica del

Inventario de Depresión de Beck, respecto a la prueba de referencia. Para lo cual se empleó el tratamiento estadístico referido al acuerdo entre-examinadores, por lo que se presentan los resultados comparativos del índice Kappa entre el BDI-II y la Escala de Depresión de Zung respecto al juicio clínico determinado por los especialistas como Gold Standard, los resultados se presentan a continuación.

Tabla 34
Parámetros de Reproductibilidad Diagnóstica (EA/BDI-II).

BDI-II	Valor	Error estándar asintótica	IC (95,0%)	Significación aproximada
Índice Kappa	0,6163	0,0539	0,5106 - 0,7219	0,000

Fuente: Cadenas, H. (2019). Recurso Personal (Cálculo EPIDAT 4.1).

En la tabla 34 se describe el grado de replicación de resultados del Inventario de Depresión de Beck, el cual exhibió Buena fuerza de concordancia, con un coeficiente equivalente a 0,61 y considerando que el límite superior con su IC corresponde a 0,72 lo cual denota buena capacidad de replicación.

Tabla 35
Parámetros de Reproductibilidad Diagnóstica (EA/EAMD).

EAMD	Valor	Error estándar asintótica	IC (95,0%)	Significación aproximada
Índice Kappa	0,4623	0,0625	0,3398 - 0,5848	0,000

Fuente: Cadenas, H. (2019). Recurso Personal (Cálculo EPIDAT 4.1).

Para la tabla 35 se presenta el grado de reproductibilidad que mostró el Gold Standard, en el que se observó un valor equivalente al 0,46 lo cual revela Moderada fuerza de concordancia, demostrando su inferioridad respecto a la prueba en estudio [BDI-II] y a pesar de asumir el valor Kappa como 0,58 al límite superior el IC no logra alcanzar el valor de concordancia demostrada por el BDI-II, en tal sentido, se responde al problema en cuestión, considerando que el BDI-II demuestra parámetros significativos de reproductibilidad diagnóstica respecto al Gold Standard.

5.2. Contrastación de Hipótesis

Habiendo con antelación fundamentado la superioridad en eficacia diagnóstica del BDI-II respecto a la Escala de Depresión de Zung como Gold Standard con los parámetros de validez, confiabilidad y reproductibilidad, es posible dar marcha al tratamiento estadístico para la comprobación de hipótesis.

La estrategia para este apartado se constituyó inicialmente para determinar la eficacia diagnóstica entre dos pruebas clínicas, una reconocida como Gold Standard y la segunda como nueva prueba en estudio, por lo que se reconoció que no sería suficiente con estimar los parámetros de precisión diagnóstica referidos a la validez, confiabilidad o reproductibilidad diagnóstica, sino también se consideró la búsqueda de un criterio metodológico y estadístico unificador, con el fin de contrastar las hipótesis formuladas en principio, en el marco de lo señalado, se consideró utilizar el Análisis de Curvas ROC con base al rigor metodológico y científico que este tipo de estudios ameritan (Altman y Bossuyt, 2005; Valle y Muñoz, 2017), puesto que este tipo de análisis aporta una medida global (agrupando los parámetros anteriormente estimados en lo que se define como Características Operativas del Receptor) que refleja que tan buena es la capacidad diagnóstica (discriminativa) de una prueba (Cerdeña y Cifuentes, 2012) la cual también puede compararse con al menos 2 pruebas más, en tal sentido, al hablar de la medida global que aportan las Curvas ROC nos referimos al Área Bajo la Curva o AUC en su abreviatura en inglés.

Curvas ROC para BDI-II respecto a Gold Standard

Por lo tanto, se inicia presentando la comprobación de eficacia diagnóstica representada por el AUC que presentó la prueba en estudio [BDI-II] a través del tratamiento estadístico con el análisis de Curvas ROC para pruebas no paramétricas.

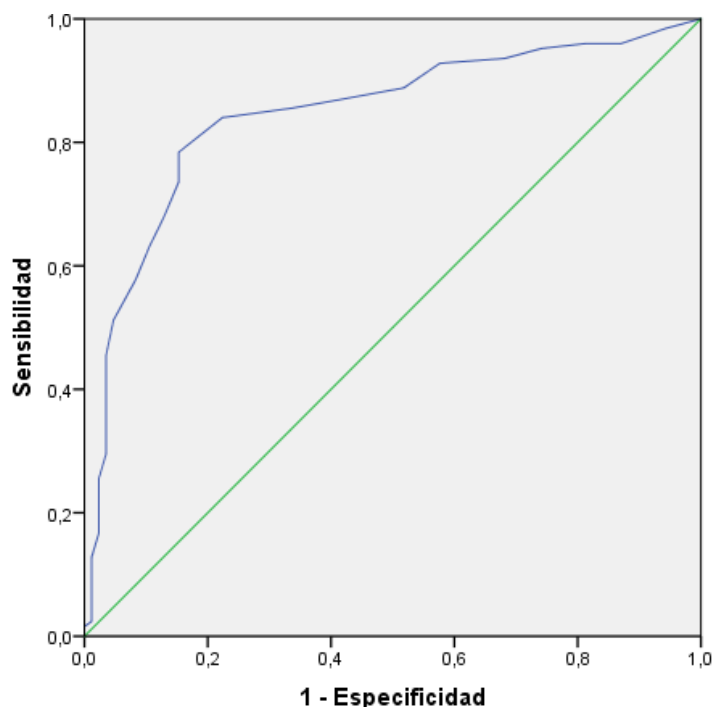


Figura 22 Curva ROC Inventario de Depresión de Beck.

Fuente: Cadenas, H. (2019). Recurso Personal (Cálculo SPSS).

Tabla 36

Área Bajo la Curva (AUC) BDI-II.

Área	Error estándar	Significación asintótica	IC: 95,0%	
			Límite inferior	Límite superior
0,846	0,028	0,000	0,791	0,901

Fuente: Cadenas, H. (2019). Recurso Personal (Cálculo SPSS).

Teniendo en cuenta los rangos para el AUC, se estableció 3 rangos posibles en función a la precisión diagnóstica que muestren las pruebas, de tipo A caracterizada por $AUC > 0,9$ categorizada como Alta Exactitud Diagnóstica, de tipo B caracterizada por $AUC > 0,8$ categorizada como Moderada y finalmente de tipo C caracterizada por $AUC > 0,7$ considera como Baja, pero a pesar de ello válida, algunos autores sugieren que valores menores a 0,7 dejarían de ser admisibles (Argimon y Jiménez, 2013; Burgueño et al. 1995; Swets, 1988). En este sentido el BDI-II demostró un rango Moderado respecto a su Exactitud Diagnóstica (además de presentar un valor de $p < 0,05$) considerando como Gold Standard a la Escala de

Depresión de Zung, cabe resaltar que llevándolo a su límite inferior por el IC aún se mantiene en un rango admisible y considerando el límite superior alcanza Alta Exactitud Diagnóstica.

Curvas ROC Comparativas BDI-II y EAMD

Habiendo demostrado que el BDI-II se considera una prueba con rangos admisibles en función al AUC que exhibió, fue posible realizar la comparación definitiva respecto a su *eficacia diagnóstica*, sin antes indicar las hipótesis formuladas para consecutivamente contrastarlas por el resultado obtenido con el análisis comparativo de Curvas ROC y su medida global [AUC]; tratamiento estadístico dispuesto al protocolo basado en evidencias que fundamenta dicho proceder al utilizar Curvas ROC (Valle & Muñoz, 2017).

Hipótesis Nula:

H₀: El Inventario de Depresión (BDI-II) de Beck no demostrará parámetros significativos de precisión diagnóstica [$AUC \leq AUC(EAMD)$] para la discriminación clínica entre pacientes adultos ($\geq 18a$) que presenten depresión de aquellos que no lo presenten, respecto a la Escala Zung de Depresión (EAMD) utilizada como Estándar de Oro.

Hipótesis Alternativa:

H_a: El Inventario de Depresión (BDI-II) de Beck demostrará parámetros significativos de precisión diagnóstica [$AUC > AUC(EAMD)$] para la discriminación clínica entre pacientes adultos ($\geq 18a$) que presenten depresión de aquellos que no lo presenten, respecto a la Escala Zung de Depresión (EAMD) utilizada como Estándar de Oro.

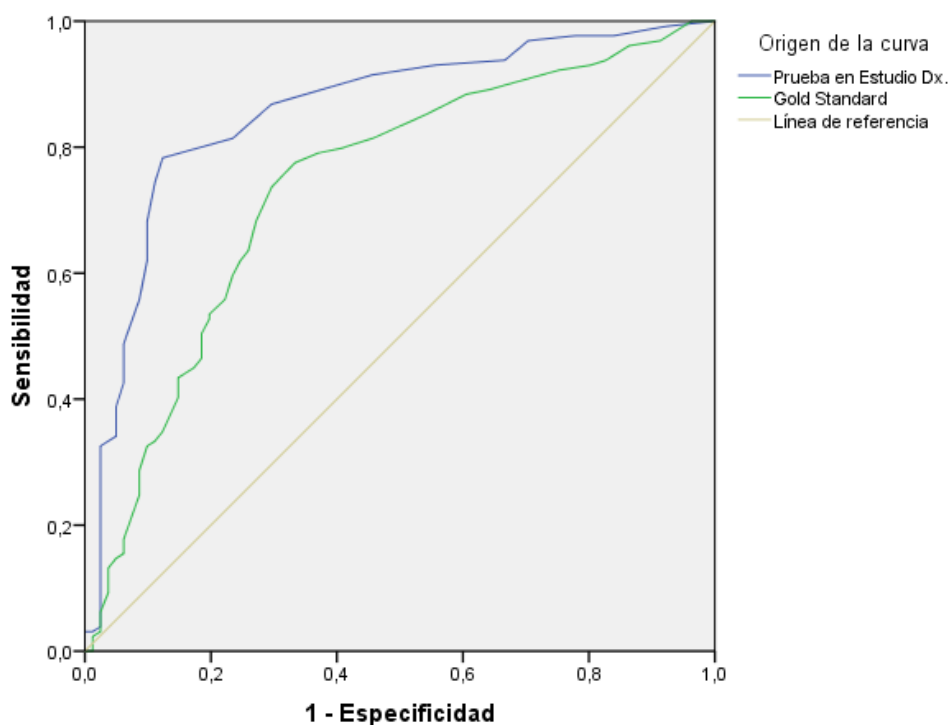


Figura 23 Curvas ROC-Comparación de Eficacia Diagnóstica.

Fuente: Cadenas, H. (2019). Recurso Personal (Cálculo SPSS).

Tabla 37

Comparación de Eficacia Diagnóstica (AUC).

Variables de resultado de prueba	Área	Error estándar	Significación asintótica	IC: 95.0%	
				Límite inferior	Límite superior
Prueba en Estudio Dx.	0,858	0,028	0,000	0,803	0,912
Gold Standard	0,739	0,036	0,000	0,668	0,809

Fuente: Cadenas, H. (2019). Recurso Personal (Cálculo SPSS).

La figura 23 muestra el Área Bajo la Curva para cada prueba al compararlas respecto al juicio clínico de especialistas como patrón de referencia, fue posible observar rápidamente como la nueva prueba en estudio (BDI-II) demostró superioridad en sus rangos de exactitud diagnóstica puesto que se superpone a la curva verde correspondiente al Gold Standard, representación que se vio corroborada por los valores tabulados en la tabla 37 en la que se obtuvo un valor Moderado [AUC=0,85] para el BDI-II, en contrastaste a un valor Bajo para el Gold Standard [AUC=0,73] y no admisible al considerar el rango al límite inferior del demarcado por el IC (con niveles de significancia menores al 0,5/ $p < 0,0005$), en este sentido y

por lo expuesto, se rechaza la hipótesis nula y en efecto, se considera que el Inventario de Depresión de Beck, demuestra parámetros significativos (superiores) de precisión diagnóstica [$AUC > AUC(EAMD)$] respecto al Gold Standard, aceptando la hipótesis alterna de superioridad en eficacia diagnóstica para la discriminación de pacientes adultos ($\geq 18a$) con o sin depresión.

Análisis Estadístico Complementario

Habiendo comprobado la hipótesis se hizo necesario complementar el análisis estadístico y dar consistencia a la significancia estadística (valor de p) demostrada por los resultados en la comparación de Curvas ROC, en ese sentido, se empleó la prueba de Rangos Signados de Wilcoxon, para determinar si existió diferencias estadísticamente significativas (puesto que es una prueba para comparar dos muestras apareadas) entre las medidas de ambas pruebas en estudio [EAMD/BDI-II], la cual es una prueba estadística no paramétrica para muestras apareadas, en la que la escala de medición puede ser ordinal, cuantitativa continua o discreta que no cuenten con distribución normal.

El procedimiento implicó inicialmente comprobar si las medidas no mantienen una distribución normal, se comenzó, por tanto, hallando la diferencia (ya que es el resultado que se analizará) entre las medidas (desarrollados en la base de datos SPSS) de cada prueba para poder estimar seguidamente la Media, Desviación Estándar y Varianza que ofrece una y otra prueba, los resultados se muestran en la siguiente tabla.

Tabla 38
Media-DE-Varianza (EAMD/BDI-II).

	Gold Standard	Prueba en Estudio Dx.	Diferencia
Media	38,94	13,44	25,5000
Desviación estándar	13,203	9,056	11,84115
Varianza	174,317	82,008	140,213

Fuente: Cadenas, H. (2019). Recurso Personal (Cálculo SPSS).

Una vez realizado ello, se analizó si esos valores presentaban normalidad en su distribución, por lo que para el contraste de normalidad se utilizó la prueba no paramétrica para una muestra en estudio, Kolmogorov-Smirnov, el cual además de demostrar si existe o no distribución normal se fundamenta en la significancia estadística, los resultados se presentan a continuación.

Tabla 39
Prueba de Kolmogorov-Smirnov (Contraste de Normalidad).

	Gold Standard	Prueba en Estudio Dx.	Diferencia	
N	210	210	210	
Parámetros normales	Media	38,94	13,44	25,5000
	Desviación estándar	13,203	9,056	11,84115
Z de Kolmogorov-Smirnov (K-S)	,127	,119	,126	
Sig. asintótica (bilateral)	,000	,000	,000	

Fuente: Cadenas, H. (2019). Recurso Personal (Cálculo SPSS).

De la tabla se pudo inferir que el valor de Z con la prueba K-S [0,12] es significativamente elevado debido al grado de significancia estadística que demuestra, la cual fue equivalente a un valor de $p < 0,0005$ ($p = 0,000$ equivale a $p < 0,0005$, si $p > 0,0005$ SPSS arroja un valor $p = 0,001$), por lo que se declara como existente la diferencia significativa de los datos con la de una distribución normal, en tal sentido, las mediciones no siguieron una distribución normal. Por lo que al no seguir una distribución normal la prueba adecuada para corroborar diferencias estadísticamente significativas en dos muestras apareadas correspondería a una prueba no paramétrica, por lo que como se expuso, se utilizó los Rangos Signados de Wilcoxon para su tratamiento estadístico, el resultado de dicho análisis se describe a continuación.

Tabla 40
Prueba de Rangos de Wilcoxon.

	Prueba en Estudio Dx. - Gold Standard
Z	-12,540
Sig. asintótica	,000

Fuente: Cadenas, H. (2019). Recurso Personal (Cálculo SPSS).

En la tabla 40 se observó que, tras el análisis estadístico, el valor de Z [-12,5] para los Rangos Signados de Wilcoxon demostró ser significativamente elevado para determinar la existencia de diferencias estadísticamente significativas entre las medidas realizadas de cada instrumento, de la cual se pudo determinar la eficacia diagnóstica [EAMD/BDI-II] y que sus resultados no se produjeron simplemente por azar en medida al nivel de significancia estadística que reportó, la cual fue equivalente a un valor de $p < 0,0005$. De tal manera se pudo aportar consistencia para la confirmación de la hipótesis alterna que se formuló en la presente, en la que se demostró superioridad de eficacia diagnóstica por parte del Inventario de Depresión de Beck sobre la Escala de Depresión de Zung siendo esta última utilizada como Gold Standard.

CAPÍTULO VI

ANÁLISIS Y DISCUSIÓN DE RESULTADOS

La presente sección se encuentra orientada a ofrecer la interpretación de los resultados por medio del análisis y la discusión de los mismos, en ese sentido de las cosas, se encuentra estructurado sistemáticamente, por lo que inicialmente se analizarán y discutirán los resultados correspondientes a los problemas específicos (en orden de su formulación) en relación a los objetivos trazados, seguido de los resultados obtenidos al contrastar la hipótesis, seguidamente el planteamiento de la discusión de consecuencias y la descripción de su proyección futura de investigaciones, para finalmente terminar dilucidando la incertidumbre entre lo estadísticamente significativo o clínicamente relevante que pudo resultar el presente estudio científico relacionado con la eficacia diagnóstica.

Análisis y discusión de Resultados Específicos

Cabe resaltar que se tuvo en cuenta los Intervalos de Confianza (límite inferior y superior) para realizar el análisis y discusión de resultados, puesto que además de estar relacionado al error aleatorio posible en la investigación, demuestra la probabilidad de que el valor real de los parámetros que se determinaron en la presente [Se, Sp, VPP, VPN, etc.] se encuentren sujetos al margen de error del 5%.

Ahora bien, cabe resaltar que los resultados expuestos a continuación son producto de la consecución efectiva del objetivo general que se planteó en la presente, la cual implicó *determinar si el BDI-II demostraría parámetros significativos (superiores) de precisión diagnóstica que permitan discriminar a pacientes adultos que presenten depresión de aquellos que no la presenten y todo ello respecto a la precisión diagnóstica que la EAMD demuestra siendo considerada Gold Standard*; en este sentido, el objetivo general fue absuelto dando paso a la obtención de parámetros epidemiológicos de ambas pruebas con las cuales se pudo contrastar la hipótesis.

Es así que, el objetivo específico inicial se orientó a *estimar los parámetros estadísticos de validez diagnóstica que demuestre el BDI-II respecto a los parámetros de la EAMD*, por lo que la propuesta inicial de la presente, se manejó en torno a si los *parámetros de validez diagnóstica* del Inventario de Depresión de Beck (BDI-II) se mostrarían significativos respecto a las de la Escala de Depresión de Zung (EAMD) catalogada como Gold Standard, por lo que de entrada, el Gold Standard (GS) debería haber demostrado superioridad sobre la nueva prueba en cuestión (puesto que se la considera como el criterio de verdad), sin embargo, el BDI-II demostró mejores parámetros de validez diagnóstica enmarcadas esencialmente en 4 parámetros diagnósticos (Sensibilidad, Especificidad, Índice de Validez e Índice de Youden) relacionados con la validez y la Prevalencia como proporción resultante de las mismas, es en este sentido que se analizan los primeros resultados, en los que el BDI-II demuestra mejor sensibilidad para detectar casos positivos cuando realmente esos casos están afectados por la condición clínica, con una proporción de éxito del 85.8% sobre 69.1% para el GS, lo que quiere decir que el BDI-II identificó como positivos exitosamente entre el 77 al 93% (de acuerdo a sus intervalos de confianza, los cuales controlan el error aleatorio) de la población que

realmente padecía la condición clínica a diferencia del GS el cual ofreció una baja proporción de éxito en medida a sus IC [58.4%-79.8%], superioridad que se ve contrastada por los hallazgos de Beltrán y colaboradores el 2012 en la que el BDI-II demuestra un 92% de Se en contraparte a lo hallado por Campo et al. el 2005 en la que la EAMD arroja un 94.7% de Se, encontrándose a la EAMD respecto a estos valores como uno de mejor Sensibilidad diagnóstica sobre el BDI, para lo cual cabe precisar que ambos valores que los instrumentos ofrecieron fueron adquiridos sobre poblaciones no clínicas, aspecto a resaltar teniendo en cuenta que ambas pruebas están destinadas al ámbito clínico de la evaluación (Beck, 2009; Lezama, 2012); demostrando que el BDI-II en la presente, es superior en Sensibilidad diagnóstica a la EAMD, aspecto contradictorio al hablar de la especificidad, puesto que respecto a este parámetro que estima la proporción de casos con resultado negativo en la población no afectada por la condición clínica, se obtuvieron proporciones de éxito cercanas, equivalentes al 77.6% para el BDI-II y 77.5% para el GS, lo cual denotaría que ambas pruebas son buenas para detectar como negativos en casos en los que la condición clínica esté ausente efectivamente, valores que contradicen lo encontrado por Beltrán, Freyre y Hernández el 2012 en su estudio de validación del BDI-II en el cual este último exhibe un 82% de Especificidad diagnóstica, lo cual implicaría que para la presente el valor es inferior respecto a dicho estudio, a diferencia de lo que sucede con la EAMD, ya que Campo y colaboradores el 2005 hallaron la Sp de 67%, por lo cual en el presente estudio la Sp sería superior a dicho hallazgo; este último pudiendo haber dependido de la población en la que se halló, ya que Campo realizó el estudio en una población universitaria no clínica, determinando de todos modos para la presente que el BDI es un instrumento con buena capacidad diagnóstica, sin embargo, al tratarse ambos de instrumentos diseñados para detectar sintomatología asociada a la depresión con escala de severidad, lo que importaría aquí sería el grado de Sensibilidad más que el Especificidad, impacto que se corroboró con la estimación de prevalencia, puesto que con el BDI-II se asume una prevalencia

(proporción de afectados en la población) equivalente al 40.4% [IC:33.6%-47.35%] en contraste al 38.5% [IC:31.7%-45.39%], lo que a escalas mayores observadas en la práctica clínica cotidiana implicaría porcentajes altamente significativos de variación, lo cual en contexto podría guiar sesgadamente a tomar decisiones clínicas sobre la población en cuestión. Finalmente, se observa que el BDI-II demuestra un Índice de validez global superior al GS [80.9% vs 74.2%], así como un Índice de Youden óptimo por estar más cercano al valor de 1 [0.63 vs 0.47], siendo incluidos estos dos últimos parámetros como nuevos a la caracterización de eficacia para ambas pruebas y pese a ello, se confirma que el BDI-II es una prueba con mejor validez diagnóstica, por demostrar parámetros de validez diagnóstica significativos y superiores a la Escala de Depresión de Zung utilizada como Gold Standard.

El segundo objetivo específico mantuvo la intención de *estimar los parámetros estadísticos de confiabilidad diagnóstica que resulte del BDI-II respecto a los parámetros de la EAMD, es así que* en lo que respecta a la siguiente propuesta relacionada a si los *parámetros de confiabilidad diagnóstica* del BDI-II se mostrarían superiores al del GS, nuevamente se confirmó que el BDI-II demuestra mejor confiabilidad diagnóstica respecto a sus 2 parámetros resultantes (Valores Predictivos y Razones de Verosimilitud), por lo que el control de incertidumbre relacionado a la probabilidad de que un paciente que diera positivo en la prueba, vaya a ser efectivamente un positivo (afectado por la condición) para el BDI-II fue superior al del GS puesto que el Valor Predictivo Positivo (desde ahora VPP) de éxito fue del 72.2% [IC:] respecto a un 65.8% [IC:] sin considerar sus IC, ya que de este modo la superioridad se afianza radicalmente, hecho equivalente para el VPN (valor predictivo negativo) en el que el BDI-II demostró mejor éxito para controlar la incertidumbre relacionada a la probabilidad de que un paciente habiendo dado negativo a la prueba vaya a ser efectivamente un caso no afectado por la condición [88.9% vs 80], lo que en resumidas cuentas se condensaría afirmando que el BDI-

II controla en un 72.2% la probabilidad de que en un caso habiendo dado un resultado positivo, vaya a ser efectivamente un caso afectado y, en un 88.9% la consideración diametralmente opuesta, parámetros que distan a los hallados por Campo y colaboradores el 2005 en los que la EAMD demuestra un 37.5% para el VPP y 98.4% respecto al VPN, siendo el primero menor al hallado por la presente y mayor al hablar del VPN, aspecto posiblemente suscitado de acuerdo a la población de la que se obtuvo los valores, la cual no fue clínica, de todos modos, los parámetros de confiabilidad diagnóstica son superiores por parte del BDI-II, no siendo contrastado con hallazgos previos puesto que a que no se registró en estudios antecedentes. Por otro lado, las Razones de Verosimilitud responden a la pregunta de cuántas veces más probable es que la prueba diagnóstica sea correcta en el grupo correcto que del incorrecto, por lo que el BDI-II demuestra del mismo modo ser una prueba más confiable ya que ofrece Razones de Verosimilitud Positivos y Negativos en función a lo señalado por Escrig-Sos y otros en el 2006, en el que dejan claro que “un buen test debe tener una RV- cercana a 0 y una RV+ alta”, sin especificar un límite superior para la RV+ pero si alejada del valor de cero, [RV+ (3.83 vs 3.08)/RV- (0.18 vs 0.40)], por lo que habría mejor probabilidad de que el BDI-II otorgara positivos a la población afectada que otorgarla a la población no afectada por la condición clínica y viceversa tratándose de la RV-, parámetros que no se llevaron a discusión puesto que no se registraron estos valores como parte de estudios antecedentes. Dejando demostrado que presenta mejores parámetros de confiabilidad (pues al tratarse de probabilidades a posteriori a la evaluación nos habla de la estabilidad de la respuesta emitida por la prueba) respecto a la del Gold Standard.

Finalmente, el tercer objetivo específico implicó *estimar los parámetros estadísticos de reproductibilidad diagnóstica que el BDI-II pudo aportar respecto a los parámetros de la EAMD*, por lo que esta última propuesta estuvo relacionada a si los *parámetros de*

reproductibilidad diagnóstica del BDI-II se mostrarían superiores al de la Escala de Depresión de Zung [EAMD], teniendo en cuenta al juicio clínico de los especialistas como Gold Standard, elección sugerida por metodología de investigación (Tornimbeni, 2008), demostrando al BDI-II como una prueba con mejor grado de replicación respecto al EAMD, con valores equivalentes al 0,61 y 0,46 para cada prueba, los cuales por criterio de Tornimbeni, Pérez y Olaz determinados el 2008, corresponden con fuerzas de concordancia Buena [0,61] y Moderada [0,46], lo cual ya demostraría mejor capacidad de replicación por parte del BDI-II, sin embargo, al considerar los intervalos de confianza, la situación parece consolidarse ya que al límite inferior del IC el BDI-II mantiene una fuerza de concordancia Moderada [0,51] y el EAMD una fuerza Débil [0,33], valores de superioridad que coincide con los hallazgos de Sanz y García el 2013 en la que el BDI-II exhibe un índice Kappa equivalente a 0,45 en una muestra clínica, respecto al hallazgo de Campo y colaboradores el 2006 en el que la EAMD exhibe un índice Kappa de 0,41 en una muestra no clínica, aspecto que podría reducir de cierto modo su desempeño en lo que reproductibilidad diagnóstica respecta, ya que el contexto idóneo para la cual ambas pruebas deberían ser reproducibles corresponde a un contexto clínico.

Análisis y discusión de Resultados: Prueba de Hipótesis

Inicialmente se corroboró si el BDI-II sería un instrumento que demuestre precisión diagnóstica teniendo en cuenta como patrón de referencia a la Escala de Depresión de Zung, por lo que el análisis de Curva ROC arrojó un área bajo la curva equivalente a 0,84 [IC: 0,79-0,90] con nivel de significación menor al 0,05% [$p=0,0005$], por lo que se cataloga como un instrumento de Moderada exactitud o precisión diagnóstica en consideración a los rangos establecidos por Swets en 1988, ratificados en 1995 por Burgueño et al. y en el 2013 por Argimon y Jiménez, los cuales son utilizados hasta el momento en diversos estudios clínicos o médicos y, teniendo en cuenta que el AUC es un indicador de eficacia general para pruebas

diagnósticas, implicaría que al utilizar la prueba en la práctica clínica sus resultados para la detección de casos positivos y el descarte de casos negativos no se produciría puramente por azar, sino que estaría atribuido a sus buenos parámetros de eficacia para el diagnóstico; medida global cercana a la hallada por Sanz y García el 2013 en el que el AUC del BDI fue de 0,83 para un estudio de rendimiento diagnóstico en una muestra también clínica, lo cual corroboraría la exactitud diagnóstica moderada (sin considerar límite superior de IC para la presente) para el BDI-II.

Teniendo en cuenta lo anterior, se dio marcha a la comparación de eficacia diagnóstica entre el BDI-II y la EAMD gracias al análisis de Curvas ROC (en la que se asumió como patrón de referencia al juicio clínico de especialistas), de los que también se obtuvo áreas bajo la curva, permitiendo dar cuenta del aporte Clínicamente Relevante de la presente investigación, puesto que al considerar que el BDI-II ofrece un AUC equivalente al 0,85 sobre el valor de 0,73 (a diferencia de lo encontrado por Campo y Díaz el 2006 con AUC de 0,89 en una muestra no clínica) de la EAMD implica no sólo un resultado más eficiente sino y sobre todo en la práctica clínica real, una mejor toma de decisiones para diversas actividades clínicas (cribados, estudios epidemiológicos, evaluaciones y actividades preventivas, etcétera) por la elección de un instrumento eficaz, desde incluso antes de evaluar al paciente hasta el impacto terapéutico que puede generar todo ello para el paciente a posteriori, fundamentando de tal modo una práctica en salud basada en la mejor evidencia científica, amplificando de tal modo la calidad de atención brindada a la población consultante, lo cual conllevaría en función a las áreas bajo la curva llevadas al límite inferior de sus Intervalos de Confianza [BDI-II=(IC:0,80-0,91)/EAMD=(IC:0,66-0,80)] a cambiar el criterio de Gold Standard y reasignarlo al del BDI-II por demostrar diferencia en su eficacia que demarca lo Clínicamente Relevante.

Discusión de consecuencias y proyección futura de investigaciones

A partir de los resultados hallados fue posible definir futuros alcances, producto del impacto generado por la nueva prueba en estudio (BDI-II), la cual se mostró como un instrumento eficaz para la detección clínica (instrumento con parámetros de validez, confiabilidad y reproductibilidad diagnóstica ideales estadísticamente) al hablar de depresión, respecto al Gold Standard, instrumento que a la fecha corresponde a la Escala de Depresión de Zung, instrumento comúnmente utilizado en la práctica clínica de los dos hospitales más representativos y de mayor complejidad resolutive en la provincia de Huancayo, los cuales además se hacen cargo de la principal carga de enfermedad en la provincia, a señalar el Hospital RDCQ Daniel Alcides Carrión y el Hospital Nacional Ramiro Priale Priale; por tanto, al tratarse de un estudio de eficacia enfocada en la precisión diagnóstica, los resultados hallados traen consigo consecuencias importantes para la Psicología clínica y de la salud, que se describen a continuación.

La posible consecuencia inmediata que los resultados de la presente puede conllevar, implicarían en esencia la integración de procedimientos clínicos asociados al diagnóstico basados en evidencias de eficacia para la práctica clínica habitual a nivel de la provincia de Huancayo (específicamente para ambos hospitales previamente señalados), por otro parte, al implicar un cambio de paradigma, podría suceder que muchos profesionales de la salud no reciban a bien tales cambios radicales, por lo que haría falta sensibilizar y comprometer a la población en general implicada en ejercer la Psicología a llevar una formación científica y crítica del aprendizaje, aunado por supuesto a la actividad científica que determine la eficacia de métodos, procedimientos, tratamientos y para el caso en específico instrumentos diagnósticos como herramientas de las que se vale el psicólogo clínico o de la salud. Por otro lado, es probable que gracias al presente estudio se comience a establecer como lo señaló

Farreras en el 2012, la jerarquía de pruebas complementarias, en la que se categoriza a los instrumentos clínicos en medida de su desempeño (eficacia) diagnóstico, basado en los mejores parámetros estadísticos y epidemiológicos que las pruebas exhiben producto de la investigación clínica.

Respecto a la proyección futura de investigaciones que demarca la presente, radican esencialmente 2 tipos de estudios posibles a realizar a posteriori, ya que teniendo claro parámetros de precisión diagnóstica se puede consecuentemente elaborar estudios del impacto terapéutico de pruebas diagnósticas como lo señala Gonzalo el 2013, en este caso del posible impacto terapéutico del BDI-II al ser administrado a casos puntuales, hablamos en este sentido del impacto sobre la toma de decisiones post-prueba para la formulación del plan terapéutico, la probabilidad de integrarla o no e incluso sobre el pronóstico del paciente; por otro lado, se podría también realizar estudios más avanzados, respecto a un diseño analítico o experimental para el estudio de pruebas diagnósticas, en el primer caso para identificar las relaciones causales que generan eficacia diagnóstica y en el segundo caso para evaluar los beneficios y riesgos asociados al uso de la prueba, todo ello bajo los términos señalados por Argimon y Jiménez el 2013.

Análisis y discusión de Resultados: Estadísticamente Significativo o Clínicamente Relevante

Finalmente, el análisis permitió dar cuenta de que los resultados hallados por la presente, son además de estadísticamente significativos, clínicamente relevantes. Lo primero, puesto que la diferencia en la eficacia diagnóstica no se produjo producto del azar, sino producto de la capacidad discriminativa que cada instrumento demostró y clínicamente relevante porque dicha diferencia de eficacia genera un impacto favorable sobre estado real del paciente.

CONCLUSIONES

1. El Inventario de Depresión de Beck en su segunda versión, demuestra parámetros significativos de validez diagnóstica, respecto a la Escala de Depresión de Zung, pudiendo ser considerada de tal modo como el nuevo patrón de referencia válido [Se:85.8%/Sp:77.6%/IV:80.9%/IY:0.63] para la evaluación de Depresión.
2. Los parámetros predictivos de confiabilidad diagnóstica que ofrece el Inventario de Depresión de Beck, demostraron ser significativos, en contraste a los ofrecidos por la Escala de Depresión de Zung, posicionándola en tal sentido como un patrón de referencia confiable [VPP:72.2%/VPN:88.9%/RV+:3.83/RV-:0.18] en la evaluación de Depresión.
3. La concordancia inter-examinadores, demostró parámetros de reproductibilidad diagnóstica del Inventario de Depresión de Beck con mayor fuerza de replicación [K:0,61/IC:0,51-0,72], respecto a la Escala de Depresión de Zung, por lo que su capacidad se homologaría a la formulación diagnóstica realizada por los especialistas.
4. Al demostrar superioridad independiente en los parámetros de validez, confiabilidad y reproductibilidad diagnóstica, el Inventario de Depresión de Beck demuestra a su vez mejor precisión respecto a la Escala de Depresión Zung gracias al área bajo la curva [AUC=0,85/IC:0,80-0,91] y las diferencias estadísticamente significativas [Z= -12,54/ $p < 0,0005$] halladas respecto a la Escala de Depresión de Zung, por lo que lo ubica como un nuevo instrumento de mejor eficacia para la detección clínica de la Depresión.

5. Por lo expuesto, se rechaza la hipótesis nula y en efecto, se considera que el Inventario de Depresión de Beck, demuestra parámetros significativos (superiores) de precisión diagnóstica [$AUC(BDI-II:0,85) > AUC(EAMD:0,73)$] respecto al Gold Standard, aceptando la hipótesis alterna de superioridad en eficacia diagnóstica para la discriminación de pacientes adultos ($\geq 18a$) con o sin depresión.

RECOMENDACIONES

1. El siguiente paso metodológico posible al presente estudio implicaría diseñar un estudio que determine el impacto terapéutico que tendría el Inventario de Depresión de Beck como prueba diagnóstica eficaz sobre los determinantes de salud a los que todos los pacientes están expuestos, por lo que se recomienda a los profesionales del área clínica interesarse por lo expuesto.
2. Teniendo en consideración que el BDI-II se ubica como una prueba eficaz para la detección clínica de la Depresión, se hace conveniente formular nuevos estudios en los que se compare su rendimiento con otras escalas o inventarios más actualizados.
3. Sólo en la práctica clínica se utilizan con mucha frecuencia una innumerable cantidad de pruebas complementarias al diagnóstico, la revisión de la literatura actual, advierte que ni el 10% de las pruebas utilizadas poseen parámetros de precisión, por lo que se recomienda a los estudiantes y profesionales orientar esfuerzos para establecer más estudios como la presente.
4. Es posible constatar la producción continua de pruebas psicológicas, sin embargo, no se cumple la misma condición al evaluar el trabajo por determinar la eficacia el rendimiento diagnóstico en el caso de las pruebas clínicas, por lo que es recomendable considerar la detención en la producción de pruebas y comenzar a elaborar jerarquías para el uso de pruebas complementarias basadas en la mejor evidencia.
5. Finalmente, se hace necesario sugerir el mantenimiento de evaluaciones completas y rigurosas para toda la instrumentación en nuestra ciencia, con la finalidad de garantizar que sólo las pruebas diagnósticas precisas sean utilizadas en nuestra práctica clínica.

REFERENCIAS BIBLIOGRÁFICAS

- Alarcón, R., Mazzotti, G., & Nicolini, H. (2005). *Psiquiatría* (2ª ed.). Washington D.C.: Manual Moderno.
- Alvarado, C. (2005). *Epistemología* (1ª ed.). Lima: Editorial Mantaro.
- Anza, I., López, A., González, F., et al. *Metodología de la Investigación y Práctica Clínica basada en la Evidencia* (1ª ed.). Murcia: Quaderna Editorial.
- Arceo, J., Ornelas, J., & Domínguez, S. (2010). *Manual de medicina basada en evidencias* (1ª ed.). México: Manual Moderno.
- Argimon, J., & Jiménez, J. (2013). *Métodos de Investigación Clínica y Epidemiológica* (4ª ed.). Barcelona: Elsevier.
- Bartrés, D., Redolar, D., Clemente, I., et al. (2008). *Bases genéticas de la conducta* (1ª ed.). Barcelona: Editorial UOC.
- Beck, A., Rush, J., Shaw, B., et al. (2010). *Terapia cognitiva para la depresión* (19ª ed.). España: Descleé de Brower.
- Beck, A., Steer, R., & Brown, G. (2009). *Inventario de Depresión de Beck (BDI-II)* (2ª ver.). Buenos Aires: Paidós.
- Caballo, V. (2002). *Manual para el tratamiento cognitivo conductual de los trastornos psicológicos* (2ª ed.). Madrid: Siglo XXI.
- Cadenas, H. (2017). *Guía de práctica clínica basada en evidencia Trastornos del estado de ánimo* (1ª ed.). Perú: Biblioteca Personal.
- Caivano, J. (1995). *Guía para realizar, escribir y publicar trabajos de investigación*. Buenos Aires: ARQUIM.
- Cardelle, F., Lorenzo, T., Baras, L., et al. (2009). *Manual del residente en Psiquiatría* (Tomo 1). Madrid: ENE Life Publicidad S.A.

- Chinchilla, A., Correas, J., Quintero, F., et al. (2011). *Breviario de urgencias psiquiátricas* (1ª ed.). España: Elsevier Masson.
- Cortese, E. (2004). *Manual de Psicopatología y Psiquiatría* (2ª ed.). Argentina: Nobuko.
- Dupont, M. (2005). *Manual de Psiquiatría* (2ª ed.). México: Editorial Alfil.
- Elkin, D. (2000). *Psiquiatría Clínica* (1ª ed.). México: McGraw-Hill.
- Farreras, P., & Rozman, C. (2012). *Medicina Interna* (17ª ed.). Barcelona: Elsevier.
- First, M., Frances, A., & Pincus, H. (2006). *DSM-IV-TR Manual de diagnóstico diferencial* (Reimpresión). Barcelona: Masson.
- Fletcher, R., & Fletcher, S. (2007). *Epidemiología Clínica* (4ª ed.). España: Lippincott Williams & Wilkins.
- Gordillo, A., Medina, Ú., & Pierdant, M. (2012). *Manual de Investigación Clínica* (1ª ed.). México: Manual Moderno.
- Green, D., & Swets, J. (1966). *Teoría de detección de señales y psicofísica* (1ª ed.). New York: Editorial John Wiley & Sons.
- Greenhalgh, T. (2016). *Cómo leer un artículo científico: Las bases de la medicina basada en la evidencia* (5ª ed.). London: Elsevier.
- Grupo CTO. (2011). *Manual CTO de Medicina y Cirugía* (8ª ed.). España: CTO Editorial.
- Hales, R., Yudofsky, S., & Talbott, J. (1996). *Tratado de psiquiatría* (2ª ed.). España: ANCORA. S.A.
- Harrison, T., Longo, D., Fauci, A., et al. (2012). *Principios de Medicina Interna* (18ª ed.). México: McGraw-Hill.
- Heinze, G., & Camacho, P. (2010). *Guía clínica para el manejo de la depresión* (1ª ed.). México: INPRFM.

- Henquin, R. (2013). *Epidemiología y Estadística para principiantes* (1ª ed.). Argentina: Corpus Editorial.
- Hernández, G. (2013). *Psicopatología básica* (5ª Edición). Colombia: Editorial Pontificia Universidad JAVERIANA.
- Hernández, I., Gil, A., Delgado, M., et al. (2005). *Manual de Epidemiología y Salud Pública Para licenciaturas y diplomados en ciencias de la salud* (1ª ed.). España: Editorial Médica Panamericana, S.A.
- Hernández, R., Fernández, C., & Baptista, C. (2014). *Metodología de la investigación* (6ª ed.). México: Mc Graw Hill.
- Hulley, S., Cummings, S., Browner, W., et al. (2014). *Diseño de Investigaciones Clínicas* (4ª ed.), Philadelphia: Lippincott Williams & Wilkins.
- Ibáñez, C. (1997). *Evaluación Psicológica lecciones introductorias* (1ª ed.). Lejona: Argitalpen Zerbitzua.
- Javier, F., & Crespo, M. (2012). *Psicología Clínica Basada en la Evidencia* (1ª ed.). Madrid: Ediciones Pirámide.
- Kaplan, B., & Sadock, V. (2009). *Sinopsis de Psiquiatría* (10ª Edición). New York: Editorial Rustica.
- Kessler, R., Chiu, W., Demler, O., et al. (2005). *Prevalence, severity and comorbidity of 12-Month* (Replication). USA: DNCS.
- Knottnerus, A., Van Well, C., Sackett, D., et al. (2002). *The evidence base of clinical diagnosis* (First Edition). London: BJM Books.
- Llanio, R., & Perdomo, G. (2003). *Propedéutica Clínica y Semiología Médica* (Tomo I). Cuba: Editorial Ciencias Médicas.
- Martell, C., Dimidjian, S., & Herman, R. (2013). *Activación conductual para la depresión Una guía clínica* (1ª ed.). España: Descleé de Brower.

- Martínez, R., Hernández, J., & Hernández, V. (2014). *Psicometría*. Madrid: Alianza Editorial.
- Morrison, J. (2014). *DSM-5 Guía para el diagnóstico clínico* (1ª ed.). New York: Manual Moderno.
- Pérez, M., Fernández, J., Fernández, C., et al. (2003). *Guía de tratamientos psicológicos eficaces* (Tomo I). Madrid: Pirámide.
- Pérez, M., Fernández, J., Fernández, C., et al. (2003). *Guía de tratamientos psicológicos eficaces I Adultos* (1ª Edición). España: Editorial Pirámide.
- Ruiz, M., Díaz, M., & Villalobos, A. (2012). *Manual de técnicas de intervención cognitivo conductuales* (1ª ed.). España: Descleé de Brower.
- Sánchez, H. & Reyes, C. (2009). *Metodología y diseños de la investigación científica* (3ª ed.). Lima: Universidad Ricardo Palma.
- Tajima, K., Molina, O., Mora, F., et al. (2014). *Manual de Psiquiatría* (3ª ed.). España: ENARM.
- Tornimbeni, S., Pérez, E., & Olaz, F. (2008). *Introducción a la psicometría* (1ª ed.). Argentina: Paidós.
- Valle, R. & Muñoz J. (2017). *Curvas ROC (Receiver-Operating-Characteristic) y sus aplicaciones* (1ª impresión). Sevilla: Universidad de Sevilla-Departamento de estadística e investigación operativa.
- Vallejo, J. (2011). *Introducción a la psicopatología y a la psiquiatría* (7ª ed.). España: Elsevier Masson.
- Vásquez, C., & Muñoz, M. (2002). *Entrevista Diagnóstica en Salud Mental: Adultos* (1ª ed.). España: Editorial Síntesis.
- Vélez, H., Rojas, W., Borrero, J., et al. (2010). *Fundamentos de medicina Psiquiatría* (5ª ed.). Colombia: Editorial CIB.

Ybarra, J., Orozco, L., & Valencia, A. (2015). *Intervenciones con apoyo empírico Herramientas fundamentales para el psicólogo clínico y de la salud*. México: Manual Moderno.

Artículos Científicos Estándar:

Altman, D., & Bossuyt, P. (2005). Estudios de precisión diagnóstica (STARD) y pronóstica (REMARK). *Med. Clin.* 125 (1), 49-55.

Ampudia, A., Sánchez, G., & Jiménez, F. (2017). Precisión diagnóstica del MMPI-2 con la personalidad delictiva: un análisis con la curva ROC. *Revista de Psicología.* 35 (1), 167-192.

Beltrán, M., Freyre, M., & Hernández, L. (2012). El Inventario de Depresión de Beck: su validez en población adolescente. *Revista TERAPIA PSICOLÓGICA.* 30 (1), 5-13.

Bravo, S., & Cruz, J. (2015). Estudios de exactitud diagnóstica: Herramientas para su interpretación. *Revista Chilena de Radiología.* 21 (4), 158-164.

Burgos, M., & Manterola, C. (2010). Cómo interpretar un artículo sobre pruebas diagnósticas. *Rev. Chilena de Cirugía.* 62 (3), 301-308.

Burgueño, M., García, J., & González, J. (1995). Las curvas ROC en la evaluación de las pruebas diagnósticas. *Med Clin.* 104 (7), 661-670.

Cabrera, V., Aragón, M., Terol, M., et al. (2015). La Escala de Ansiedad y Depresión Hospitalaria (HAD) en Fibromialgia: Análisis de sensibilidad y especificidad. *Revista TERAPIA PSICOLÓGICA.* 33 (3), 181-193.

Campo, A., Díaz, L., & Rueda, G. (2006). Validez de la escala breve de Zung para tamizaje del episodio depresivo mayor en la población general de Bucaramanga. *Biomédica.* 26 (4), 15-23.

- Campo, A., Díaz, L., Rueda, G., et al. (2005). Propuesta de una versión abreviada de la escala de Zung para depresión. *Colombia médica*. 36 (3), 168-172.
- Campo, A., Díaz, L., Rueda, G., et al. (2005). Validación de la escala de Zung para depresión en universitarias de Bucaramanga, Colombia. *Revista Colombiana de Psiquiatría*. 34 (1), 54-62.
- Carbonell, M. (2016). Valor diagnóstico de la Escala de Ansiedad y Depresión de Goldberg (EAD-G) en adultos cubanos. *Universitas Psychologica*. 15 (1), 15-29.
- Cerda, J., & Cifuentes, L. (2012). Uso de curvas ROC en investigación clínica. Aspectos teórico-prácticos. *Rev Chil Infect*. 29 (2), 138-141.
- De la Torre, J., Shimabukuro, R., Varela, L., et al. (2006). Validación de la versión reducida de la escala de depresión geriátrica en el consultorio externo de geriatría del Hospital Nacional Cayetano Heredia. *Act Med Per*. 23 (3), 144-147.
- Díaz, G., Reyes, H., López, A., et al. (2006). Validación de una escala clinimétrica para el diagnóstico de depresión en pacientes con diabetes mellitus tipo 2, en unidades de atención primaria. *Rev Invest Clin*. 58 (5), 432-440.
- Domínguez, E., & González, R. (2002). Análisis de las curvas ROC: Un método útil para evaluar procedimientos diagnósticos. *Rev. Cubana Endocrinol*. 13 (2), 169-176.
- Donis, J. (2012). Evaluación de la validez y confiabilidad de una prueba diagnóstica. *Avances en Biomedicina*. 1 (2), 73-81.
- Donis, J. (2013). Tipos de diseños de los estudios clínicos y epidemiológicos. *Rev Avanz Biomed*. 2 (2), 76-99.
- Echeburúa, E., & Corral, P. (2005). Reflexiones ante la formación de los psicólogos. Retos del futuro. *Análisis y Modificación de conducta*. 1 (31), 175-188.

- Echeburúa, E., Salaberría, K., Corral, P., et al. (2010). Terapias psicológicas basadas en la evidencia: limitaciones y retos del futuro. *Revista Argentina de clínica psicológica*. *19* (3), 247-256.
- Escrig Sos, J., Martínez, D., & Miralles, JM. (2006). Pruebas diagnósticas: nociones básicas para su correcta interpretación y uso. *Cir. Esp.* *79* (5), 267-273.
- Fagan, T. (1975). Nomogram for Bayes theorem. *N Engl J Med.* *293* (5), 257-262.
- Frías, M., & Pascual, J. (2003). Psicología Clínica Basada en Pruebas: Efecto del tratamiento. *Papeles del Psicólogo*. *85* (1), 11-18.
- Gómez, C., & Sánchez, R. (1998). Cálculo del tamaño de la muestra en psiquiatría y salud mental (principios básicos para su estimación). *Rev Col Psiquiatría*. *27* (2), 131-142.
- González, M., & Velarde, C. (2012). Lista de comprobación de estudios de precisión de pruebas diagnósticas: declaración STARD. *Evid. Pediatr.* *8* (43), 1-4.
- González, M., & Velarde, C. (2012). Lista de comprobación de estudios sobre pruebas diagnósticas incluidos en las revisiones sistemáticas: declaración QUADAS. *Evid. Pediatr.* *8* (20), 1-6.
- Harper, R. Reeves, B. (1999). Reporting of precision of estimates for diagnostic accuracy: a review. *BMJ.* *318* (7194), 1322-1333.
- Hernández, L., Palacio, A., & Freyre, M. (2011). La perspectiva dimensional de la psicopatología. *Revista Mexicana de Psicología*. *28* (2), 111-120.
- Ibáñez, C., & Echeburúa, E. (2015). Función y limitaciones del pronóstico en la evaluación diagnóstica en el ámbito de la psicología clínica. *Clínica y Salud*. *1* (2), 17-22.
- Landa, N., Goñi, A., García, E., et al. (2008). Concordancia en el diagnóstico entre atención primaria y salud mental. *Aten Primaria*. *40* (6), 285-289.

- Lejuez, C., Hopko, D., & Hopko, S. (2001). A brief Behavioral Activation Treatment for Depression. *BEHAVIOR MODIFICATION*. 25 (2), 255-286.
- Letelier, S., & Moore, P. (2003). Evidence Based Medicine Revisited. *Revista Médica de Chile*. 131 (8), 939-946.
- Lezama, S. (2012). Propiedades psicométricas de la escala de Zung para síntomas depresivos en población adolescente escolarizada colombiana. *PSYCHOL. AV. DISCIP.* 1 (6), 91-101.
- Loong, TW. (2003). Understanding sensitivity and specificity with the right side of the brain. *BMJ*. 327 (7417), 716-719.
- Madero, R., Pérez, E., & San José, B. (2007). Estadística para pediatras III. *An Pediatr Contin*. 5 (1), 63-67.
- Manes, J., & Ayuso, J. (2010). Sobrediagnóstico y sobretratamiento del trastorno depresivo mayor en atención primaria: un fenómeno en auge. *Aten Primaria*. 42 (1), 47-49.
- Manterola, C. (2009). Cómo interpretar un artículo sobre pruebas diagnósticas. *Rev Med Clin Condes*. 20 (5), 708-717.
- Manterola, C. (2009). Estudios observacionales. Los diseños utilizados con mayor frecuencia en investigación clínica. *Rev. Médica*. 20 (4), 539-548.
- Manterola, C., & Zavano, D. (2009). Cómo interpretar los “Niveles de evidencia” en los diferentes escenarios clínicos”. *Rev Chilena de Cirugía*. 61 (6), 582-595.
- Manterola, C., Asenjo, C., & Otzen, T. (2014). Jerarquización de la evidencia. Niveles de evidencia y grados de recomendación de uso actual. *Rev chilena Infectol*. 31 (6), 705-718.

- Martín, M., Riquelme, A., Ortigosa, J., et al. (2012). Utilidad del inventario de Depresión Rasgo-Estado (IDERE) para el diagnóstico clínico. Un estudio con muestras de 4 países iberoamericanos. *Anales de Psicología*. 28 (3), 763-771.
- Martina, M., Ara, M., Gutiérrez, C., et al. (2017). Depresión y factores asociados en la población peruana adulta mayor según la ENDES 2014-2015. *An Fac Med*. 78 (4), 393-397.
- Martínez, J., Onís, C., Dueñas, R., et al. (2002). Versión española del cuestionario de Yesavage abreviado (GDS) para el despistaje de depresión en mayores de 65 años: adaptación y validación. *MEDIFAM*. 12 (10), 620-630.
- Molina, M. (2013). Lectura crítica en pequeñas dosis Características de las pruebas diagnósticas. *Rev. Pediatr Aten Primaria*. 15 (1), 169-173.
- Montenegro, M., & Kilstein, J. (2014). Utilidad del PHQ-9 en el diagnóstico de trastorno depresivo en una sala de clínica médica. *IntraMed Journal*. 6 (1), 1-12.
- Moriana, J., & Martínez, V. (2011). La psicología basada en la evidencia y el diseño y evaluación de tratamientos psicológicos eficaces. *Revista de Psicopatología y Psicología Clínica*. 16 (2), 81-100.
- Mosqueira, N., Gutiérrez, A., Escobedo, S., et al. (2015). Costo de enfermedades mentales prevalente en Perú. *Abt Associates-HFG*. 18 (11), 812-814.
- Navarro, J., Moscoso, M., & Calderón, G. (2017). La investigación de la depresión en adolescentes en el Perú: una revisión sistemática. *Liberabit*. 23 (1), 57-74.
- Otero, W., & Sierra, F. (2003). Medidas de impacto o magnitud del beneficio de una intervención. *Rev Colomb Gastroenterol*. 18 (4), 46-48.
- Pérez, O., García, C., Grases, N., et al. (2015). Programa de soporte a la primaria: Un análisis descriptivo y de concordancia diagnóstica entre médico/a de atención primaria y psicólogo/a. *Rev. Asoc. Esp. Neuropsiq*. 35 (126), 257-266.

- Piazza, M., & Fiestas, F. (2014). Prevalencia anual de trastornos y uso de servicios de salud mental en el Perú: resultados de estudio mundial de salud mental. *Revista Perú Med Exp Salud Pública*. *31 (1)*, 30-38.
- Sánchez, M. (2001). La medicina basada en evidencias en México: ¿lujo o necesidad? *Anales Médicos del Hospital ABC*. *46 (2)*, 97-103.
- Sanz, J. (2014). Recomendaciones para la utilización de la adaptación española del Inventario de Ansiedad de Beck (BAI) en la práctica clínica. *Clínica y Salud*. *25 (1)*, 39-48.
- Sanz, J., & García, M. (2013). Rendimiento diagnóstico y estructura factorial del Inventario de Depresión de Beck-II (BDI-II). *Anales de Psicología*. *29 (1)*, 66-75.
- Sanz, J., García, M., Espinosa, R., et al. (2005). Adaptación española del Inventario de Depresión de Beck-II (BDI-II): propiedades psicométricas en pacientes con trastornos psicológicos. *Clínica y Salud*. *16 (2)*, 121-142.
- Sierra, F. (2003). La sensibilidad y especificidad: entendiendo su origen y utilidad real. *Rev. Colomb Gastroenterol*. *18 (3)*, 180-182.
- Swets, J. (1988). Measuring the Accuracy of Diagnostic Systems. *Science*. *240 (4857)*, 1285-1293.
- Vargas, B., Villamil, V., Rodríguez, C., et al. (2011). Validación de la escala Kessler 10 (K-10) en la detección de depresión y ansiedad en el primer nivel de atención. *Salud Mental*. *34 (4)*, 323-331.
- Vega, J., Coronado, O., & Mazzotti, G. (2014). Validez de una versión en español del Inventario de Depresión de Beck en pacientes hospitalizados de medicina general. *Revista Neuropsiquiatría*. *77 (2)*, 95-103.

Vera, P. (2004). Estrategias de intervención en psicología clínica: las intervenciones apoyadas en la evidencia. *Liberabit. 1 (10)*, 4-10.

Zamora, J. (2008). Una revisión del estado actual de la metodología de evaluación de pruebas diagnósticas. *Ciberesp. 1 (2)*, 3-37.

Zevallos, S. (2015). Prevalencia de depresión en poblaciones andinas. *Anales de Salud Mental. 31 (2)*, 9-22.

Zweig, M., & Campbell, G. (1993). Receiver Operating Characteristic (ROC) Plots: A fundamental evaluation tool in clinical medicine. *Clin Chem. 39 (8)*, 561-577.

Tesis de Grado:

Gonzalo, F. (2013). *Validez de los elementos diagnósticos en endometriosis. Aplicación al análisis de decisión clínica*. Tesis para optar el grado de Doctor en Medicina Humana. Universidad de Alcalá, Madrid.

Vacca, M. & Mercado, M. (2005). *Determinación de las características operativas de las pruebas serológicas con cepas colombianas de trypanosoma cruzi utilizadas para el diagnóstico de enfermedad de Chagas*. Tesis para optar el grado de Maestro en Epidemiología Clínica. Pontifica Universidad Javeriana, Bogotá.

Organizaciones Como Autores:

Badillo, L. (2017). *Manual del test Zung de Ansiedad y Depresión*. El Tambo: Testoteca Servicio de Psicología HNRPP.

Badillo, L., & Cadenas, H. (2017). *Monitoreo de Salud Mental*. El Tambo: Servicio de Psicología, Hospital Nacional Ramiro Prialé Prialé.

Cadenas, H. (2008). *Reporte de Morbilidad*. Huancayo: Oficina de Estadística, Hospital RDCQ Daniel Alcides Carrión.

Instituto Especializado de Salud Mental Honorio Delgado-Hideyo Noguchi. (2003). *Estudio Epidemiológico en Salud Mental en la Sierra Peruana*. Perú: Anales de Salud Mental.

Minsa. (2010). *Salud Mental y Derechos Humanos Supervisión de la política pública, la calidad de los servicios y la atención a poblaciones vulnerables* (2ª ed.). Perú: Biblioteca Nacional del Perú.

OMS. (2000). *Clasificación Internacional de Enfermedades* (10ª ed.). España: Editorial Médica Panamericana.

Medio Audiovisual:

Escuela Gallega de Administración Sanitaria. (García, I). (2012). *Introducción al módulo de metodología de la investigación* [Fichero de vídeo]. De:
<https://www.youtube.com/watch?v=rLSHaDOSvbk&t=6s>.

Escuela Gallega de Administración Sanitaria. (Gonzales, A). (2012). *La pregunta a investigar* [Fichero de vídeo]. De:
<https://www.youtube.com/watch?v=DJXpagArF4A&t=10s>.

Escuela Gallega de Administración Sanitaria. (Rodríguez, M). (2013). *Diseñando un estudio de pruebas diagnósticas* [Fichero de vídeo]. De:
<https://www.youtube.com/watch?v=-w6dCgzH0ks>.

Artículos de Periódico:

Aguirre, A. (2017, mayo) *Cifras de salud mental en el país*. Correo Piura. 2.

Al editor. (2015, junio). *Estado de la salud mental Perú*. Perú21. Recuperado el:
<https://peru21.pe/peru/cifras-salud-mental-peru-menores-ocupan-70-atenciones-infografia-394376>.

Medina, R. (2016, septiembre). *La salud mental de peruanos*. La República. 4.

ANEXOS

Anexo 1: Matriz de Consistencia

Formulación del Problema de Estudio	Formulación de Objetivos	Marco Teórico	Hipótesis de Estudio	Metodología de Investigación
<p>¿Demostrará el Inventario clínico de Depresión de Beck parámetros significativos de precisión diagnóstica al distinguir entre pacientes adultos que presenten depresión de los que no lo presenten, al utilizar la Escala de Depresión de Zung como Estándar de oro?</p> <p>Problemas Específicos:</p> <p>a) ¿Demuestra el inventario clínico de Depresión de Beck parámetros significativos de validez diagnóstica, respecto a la escala de Depresión de Zung utilizada como prueba de referencia?</p> <p>b) ¿Expresa el inventario clínico de Depresión de Beck parámetros significativos de confiabilidad diagnóstica, respecto a la escala de Depresión de Zung utilizada como prueba de referencia?</p>	<p>1.- General: Determinar si el Inventario clínico de Depresión de Beck (BDI-II) demuestra parámetros significativos de precisión diagnóstica que permita discriminar pacientes adultos ($\geq 18a$) que presenten depresión de aquellos que no lo presenten, respecto a la Escala de Zung (EAMD) utilizada como Estándar de Oro.</p> <p>2.- Específicos:</p> <p>a) <i>Estimar</i> parámetros estadísticos de validez diagnóstica que aporta el Inventario de Beck para la detección clínica de Depresión, respecto a la prueba de referencia.</p> <p>b) <i>Valorar</i> parámetros estadísticos de confiabilidad diagnóstica que aporta el Inventario de Beck para la detección clínica de Depresión, respecto a la prueba de referencia.</p> <p>c) <i>Valuar</i> parámetros estadísticos de reproductibilidad diagnóstica del Inventario de Depresión de Beck, respecto a la prueba de referencia.</p>	<p>1.- Marco Referencial: <u>Prueba de referencia:</u> Campo, A., Díaz, Luis., Rueda, G., Barros, J. en Colombia el año 2005 realizaron un estudio de validación de la escala de Zung, aplicado a 110 universitarias para determinar su valor diagnóstico, teniendo como gold standard a la entrevista estructurada SCID-I, encontrado parámetros de sensibilidad, 94,7% (IC 95%: 90,5-99,7); especificidad, 67,0% (IC 95%: 56,3-76,3); concluyendo que demuestra una precisión aceptable.</p> <p>El mismo equipo de investigación realizó en ese mismo año una validación de la escala abreviada del Zung, aplicada a 138 estudiantes universitarias, manteniendo la misma prueba de referencia (SCID-I), encontrando, alfa de Cronbach fue 0.845, la sensibilidad 0.92%, la especificidad 0.71%, kappa 0.383 y el área bajo de la curva ROC 0.84 para el punto de corte de 22, concluyendo que la ED-Z de 20 reactivos es un instrumento válido al igual que versión abreviada de 10 reactivos.</p> <p>El mismo equipo de investigación realizó otro estudio de validación de la escala breve de Zung el 2006, considerando una muestra mayor de 266 personas adultas, como gold standard a SCID-I, en la que la ES-Z demostró una sensibilidad de 95,5%, especificidad de 70,3%, kappa media de Cohen de 0,415 y 0,898 de área bajo la curva receptor-operador, concluyendo que la escala abreviada presenta propiedades psicométricas similares a la escala completa, considerándose como un buen instrumento de detección.</p>	<p>1.- Hipótesis General: Hipótesis Nula: <i>El Inventario de Depresión (BDI-II) de Beck no demostrará parámetros significativos de precisión diagnóstica [AUC\leqAUC(EAMD)] para la discriminación clínica entre pacientes adultos ($\geq 18a$) que presenten depresión de aquellos que no lo presenten, respecto a la Escala Zung de Depresión (EAMD) como Estándar de Oro.</i></p> <p>Hipótesis Alterna: <i>El Inventario de Depresión (BDI-II) de Beck demostrará parámetros significativos de precisión diagnóstica [AUC$>$AUC(EAMD)] para la discriminación clínica entre pacientes adultos ($\geq 18a$) que presenten depresión de aquellos que no lo presenten, respecto a la Escala Zung de Depresión (EAMD) como Estándar de Oro.</i></p> <p>2.-Hipótesis Específicas: <i>He1: El inventario clínico de Depresión de Beck demuestra parámetros significativos de validez diagnóstica, respecto a la escala de Depresión de Zung utilizada como prueba de referencia.</i></p>	<p>1.- Método: La presente se ha diseñado bajo el método científico, con metodología propia de una investigación clínica y epidemiológica.</p> <p>2.- Tipo de Investigación: Corresponde, de entre los diversos estudios clasificados esencialmente por su finalidad (es decir, descriptiva o analítica), a los Estudios de Precisión Diagnóstica de tipo descriptivo.</p> <ul style="list-style-type: none"> Desde la perspectiva convencional la presente correspondería a un tipo de investigación tecnológica, debido a que ofrece respuestas a problemas técnicos como lo son instrumentos de evaluación clínica psicológica. <p>3.- Diseño de Investigación: Su diseño corresponde: Investigación observacional, no experimental, descriptiva y transversal con enfoque prospectivo.</p> <p>4.- Población: Los manuales de Investigación Clínica y los estándares internacionales de calidad para el estudio de pruebas diagnósticas (ya sea STARD o QUADAS) sugieren una determinación muestral en función de la prevalencia estimada de la enfermedad.</p> <p><u>Se determina el tamaño muestral bajo las siguientes consideraciones:</u></p> <ul style="list-style-type: none"> Intervalo de Confianza (IC) al 95% (intervalo para el control de error alfa recomendado). Margen de error del 10% y PA del 90% (valor de PA sugerido para el control de error beta). Los parámetros de Se y Sp basados en estudios de validez diagnóstica del método de

<p>c) ¿Exhibe el inventario clínico de Depresión de Beck parámetros significativos de reproductibilidad diagnóstica, respecto a la escala de Depresión de Zung utilizada como prueba de referencia?</p>		<p>Prueba en estudio: Sanz, J. García, M. Espinosa, R. Fortún, M. Vázquez, C. en el año 2005 en España, realizaron un estudio de Adaptación española del BDI-II, a través del estudio de sus propiedades psicométricas, con una muestra de 305 pacientes, su validez discriminativa correlacionó adecuadamente y estadísticamente significativa ($p < 0,01$) con otras escalas, concluyendo que el BDI es un instrumento válido de detección y cuantificación de síntomas depresivos, aunque su utilidad como herramienta para el diagnóstico diferencial de la depresión es una cuestión pendiente de investigación.</p> <p>Beltrán, M. Freyre, M. Hernández, L. en México el año 2012 realizaron un estudio en el que exploraron la Validez del Inventario de Depresión de Beck, en una muestra de 385 adolescentes, basando su patrón de referencia a los criterios CIE-10, obteniendo una sensibilidad de 0.92% y una especificidad de 0.59%, a pesar de que estos valores no son producto del mejor corte con Curvas ROC.</p> <p>Sanz, J. y García, M. en el año 2013, en España, realizaron un estudio de Rendimiento diagnóstico y estructura factorial del BDI-II, considerando como método de referencia al SCID-I, con una muestra de 322 pacientes, obteniendo como mejor punto de corte 25, la cual proporciona un 73,5% de sensibilidad y 80,3% de especificidad, lo que supone un índice de eficacia diagnóstica y de concordancia (Kappa) de 78.9 y 0.45, lo cual supondría una capacidad discriminadora buena.</p>	<p><i>He2: El inventario clínico de Depresión de Beck <u>expresa</u> parámetros significativos de confiabilidad diagnóstica, respecto a la escala de Depresión de Zung utilizada como prueba de referencia.</i></p> <p><i>He3: El inventario clínico de Depresión de Beck <u>exhibe</u> parámetros significativos de reproductibilidad diagnóstica, respecto a la escala de Depresión de Zung utilizada como prueba de referencia.</i></p>	<p>referencia previo que demostró mejor precisión (Depresión= Se:92,0%; Sp:71,0%).</p> <ul style="list-style-type: none"> El cálculo de la prevalencia se realiza gracias a los reportes de morbilidad a los que pude tener acceso por parte de la Oficina de Estadística del Hospital RDCQ "Daniel Alcides Carrión, de los que se estima la prevalencia en el servicio de psicología y psiquiatría equivalente en: Depresión= 22,5%. <p>5.- Muestra: Respecto al tipo de muestreo se procedió con un muestreo no probabilístico, con selección de casos consecutivos.</p> <ul style="list-style-type: none"> La población de estudio estará contextualizada por pacientes que acuden al Servicio de Psicología y Psiquiatría de dicho Hospital, conformado por 156 casos en total (con la finalidad de aproximarse a la PA del 9%), sin embargo, gracias al control estadístico de las no respuestas o casos inválidos, se determinó un tamaño muestral correspondiente a 210 casos. <p>6.- Definición Operacional: Pruebas en estudio (BDI-II): El BDI-II es un instrumento con buenas propiedades psicométricas, aporta resultados válidos, confiables y reproducibles, al momento de detectar severidad sintomatológica depresiva.</p> <p>Gold standard (EAMD): La EAMD, es un instrumento que demuestra propiedades psicométricas óptimas, aporta resultados válidos, confiables y reproducibles, al detectar la severidad sintomatológica relacionada a la depresión.</p> <p>7.- Técnicas de Recolección de Datos e Instrumentos: Gold standard: La Escala de Autoevaluación de Zung para la depresión (EAMD).</p>
---	--	--	--	--

2.- Marco Conceptual:

- Estudio de Precisión Diagnóstica (EPD): Es un tipo de estudio descriptivo que requiere una mención especial. Tratan de conocer la capacidad diagnóstica para clasificar a una persona según la presencia o ausencia de enfermedad.
- Discriminación Diagnóstica: Capacidad de alguna prueba diagnóstica para diferenciar sujetos sanos versus enfermos.
- Area Under the Curve (AUC): Medida refleja qué tan bueno es el test para discriminar pacientes con y sin la enfermedad a lo largo de todo el rango de puntos de corte posibles.
- Estándar de Oro (Gold Standard): Es una prueba para determinar la presencia o ausencia de alguna enfermedad, sintomatología o entidad nosológica, la cual posee superior rendimiento diagnóstico.
- Curva ROC: Características operativamente funcionales de una prueba diagnóstica, el cual constituye un método estadístico para determinar la exactitud diagnóstica de tests.

3.- Teorías Básicas Referenciales:

- Modelo clínico psiquiátrico de la depresión.
- Modelo conductual, cognitivo y contextual de la depresión.
- Medicina basada en la evidencia.
- Psicología clínica basada en evidencia.
- Terapias psicológicas basadas en la evidencia.
- Estudios de eficacia diagnóstica.

Prueba en estudio:

La prueba que de la que se demostró su eficacia diagnóstica fue el Inventario clínico de Depresión (BDI-II) de Beck.

8.- Técnicas de Procesamiento de Datos:

La actividad científica se realizó a través del software Excel en su versión del 2016, además de haberse valido del software para procesamiento de datos estadísticos SPSS en su versión más actualizada 24.0.0, con el que se realizó varios de los análisis estadísticos presentados en la presente [Incluye el análisis de Curvas ROC para pruebas no paramétricas, así como el análisis de Rangos Signados de Wilcoxon con la que se afianzó la significación estadística para concluir que el BDI-II es una prueba con mejor eficacia diagnóstica que el Gold Standard]. Finalmente, también se utilizó el programa para el análisis epidemiológico de datos tabulados EPIDAT, con el que se halló parámetros de validez, confiabilidad y reproductibilidad diagnóstica con sus correspondientes intervalos de confianza.

Anexo 2: Matriz de Operacionalización de Instrumento

<i>Variable</i>	<i>Definición Conceptual</i>	<i>Definición Operacional(*)</i>	<i>Dimensión</i>	<i>Sub-Dimensiones</i>	<i>Categorías/Criterios de medición</i>	<i>Escala de Medición</i>
Variable Dependiente: Inventario de Depresión (BDI-II) de Beck	BDI-II: El Inventario de Depresión de Beck es un instrumento clínico, adaptado en su segunda versión a criterios diagnósticos del DSM-IV, el cual se encuentra orientado a la identificación de sintomatología depresiva y estimarla en una escala de gravedad que confiere sus intervalos a varios puntos de corte óptimos (Beck, Steer & Brown, 2009).	En el presente estudio la definición operacional se orientó a describir propiedades estadísticas de rendimiento: <i>El Inventario de Depresión de Beck es un instrumento con buenas propiedades psicométricas, puesto que aporta resultados válidos, confiables y reproducibles, al momento de detectar severidad sintomatológica depresiva, con base en sus puntos de corte operativo.</i>	Validez	Sensibilidad [92%] Especificidad [82%]	<u>Rangos determinados por sus mejores puntos de cortes para la discriminación diagnóstica.</u> (Beck, Steer & Brown, 2009). Depresión mínima (≤ 10). Depresión leve (11-17). Depresión moderada (18-29). Depresión severa ($30 \geq$).	Cualitativa ordinal/No paramétrico
			Confiabilidad	Alfa de Cronbach [0.89]		
			Reproductibilidad	Índice Kappa [0.45]		
			Eficacia	Área Bajo la Curva [0.83/como Gold Standard: SCID-I]		

<i>Variable</i>	<i>Definición Conceptual</i>	<i>Definición Operacional(*)</i>	<i>Dimensión</i>	<i>Sub-Dimensiones</i>	<i>Categorías/Criterios de medición</i>	<i>Escala de Medición</i>
Variable Independiente: Escala de Zung de Depresión (EAMD)	EAMD: La escala de autoevaluación de la depresión (EAMD) se encuentra destinada a la estimación numérica o cuantitativa de la depresión como condición clínica, en 4 rangos de severidad clínica.	En el presente estudio la definición operacional está destinada a describir propiedades estadísticas de rendimiento: <i>La Escala de Depresión de Zung, es un instrumento que demuestra propiedades psicométricas óptimas, a razón de aportar resultados válidos, confiables y reproducibles, al detectar la severidad sintomatológica relacionada a la depresión, con base en sus rangos diagnósticos.</i>	Validez	Sensibilidad [94.7%/IC:90.5%-99.7%] Especificidad [67.0%/IC:56.3%-76.3%]	<u>Rangos no fundamentados por puntos de corte para su grado de discriminación diagnóstica.</u> Depresión normal (≤ 40). Depresión mínima a moderada (41-47). Depresión marcada a severa (48-55). Depresión en grado máximo ($56 \geq$).	Cualitativa ordinal/No paramétrica.
			Confiabilidad	VPP [37.5%] VPN [98.4%] Alfa de Cronbach [0.85] Test Re-Test [0.84]		
			Reproductibilidad	Índice Kappa [0.39]		
			Eficacia	Área Bajo la Curva [0.84/como Gold Standard: SCID-I]		

Anexo 3: Instrumentos de Investigación

Inventario de Depresión de Beck

<p style="text-align: center;">BDI-II</p> <p>Nombre: Edad: Sexo: Estado civil: Profesión: Estudios:</p> <p>Instrucciones: Este cuestionario consiste en 21 grupos de afirmaciones. Por favor, lee con atención cada uno de ellos y, a continuación, señala cuál de las afirmaciones de cada grupo describe mejor el modo en el que se ha sentido DURANTE LAS DOS ÚLTIMAS SEMANAS, INCLUYENDO EL DÍA DE HOY. Rodea con un círculo el número que se encuentre escrito a la izquierda de la afirmación que haya elegido. Si dentro del mismo grupo hay más de una afirmación que considere igualmente aplicable a su caso, señálela también. Asegúrese de leer todas las afirmaciones dentro de cada grupo antes de efectuar la elección.</p>	<table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 50%; vertical-align: top;"> <p>1. Tristeza</p> <p>0 No se siente triste habitualmente. 1 Se siente triste gran parte del tiempo. 2 Se siente triste continuamente. 3 Se siente tan triste o tan desgraciado que no puede soportarlo.</p> <p>2. Pesimismo</p> <p>0 No está desanimado sobre su futuro. 1 Se siente más desanimado sobre su futuro que antes. 2 No espera que las cosas mejoren. 3 Siente que su futuro es desesperanzador y que las cosas sólo empeorarán.</p> <p>3. Sentimientos de fracaso</p> <p>0 No se siente fracasado. 1 Ha fracasado más de lo que debería. 2 Cuando mira atrás, mira fracaso tras fracaso. 3 Se siente una persona totalmente fracasado.</p> <p>4. Pérdida de placer</p> <p>0 Disfruta de las cosas que le gustan tanto como antes. 1 No disfruta de las cosas tanto como antes. 2 Obtiene muy poco placer de las cosas con las que antes disfrutaba. 3 No obtiene ningún placer de las cosas con las que antes disfrutaba.</p> <p>5. Sentimientos de culpa</p> <p>0 No se siente especialmente culpable. 1 Se siente culpable de muchas cosas que ha hecho o debería haber hecho. 2 Se siente bastante culpable la mayor parte del tiempo. 3 Se siente culpable consistentemente.</p> </td> <td style="width: 50%; vertical-align: top;"> <p>6. Sentimientos de castigo</p> <p>0 No siente que está siendo castigado. 1 Siente que puede ser castigado. 2 Espera ser castigado. 3 Siente que está siendo castigado.</p> <p>7. Inestabilidad con uno mismo</p> <p>0 Siente lo mismo que antes sobre usted mismo. 1 Ha perdido confianza en sí mismo. 2 Está desconfiado consigo mismo. 3 No se gusta.</p> <p>8. Autoexigencia</p> <p>0 No se critica o se culpa más que antes. 1 Es más crítico consigo mismo de lo que solía ser. 2 Critica todos sus defectos. 3 Se culpa por todo lo malo que sucede.</p> <p>9. Pensamientos o deseos de suicidio</p> <p>0 No tiene ningún pensamiento de suicidio. 1 Tiene pensamientos de suicidio, pero no los llevaría a cabo. 2 Le gustaría suicidarse. 3 Se suicidaría si tuviese la oportunidad.</p> <p>10. Llanto</p> <p>0 No llora más de lo que solía hacerlo. 1 Lloro más de lo que solía hacerlo. 2 Lloro por cualquier cosa. 3 Tiene ganas de llorar continuamente, pero no puede.</p> </td> </tr> </table>	<p>1. Tristeza</p> <p>0 No se siente triste habitualmente. 1 Se siente triste gran parte del tiempo. 2 Se siente triste continuamente. 3 Se siente tan triste o tan desgraciado que no puede soportarlo.</p> <p>2. Pesimismo</p> <p>0 No está desanimado sobre su futuro. 1 Se siente más desanimado sobre su futuro que antes. 2 No espera que las cosas mejoren. 3 Siente que su futuro es desesperanzador y que las cosas sólo empeorarán.</p> <p>3. Sentimientos de fracaso</p> <p>0 No se siente fracasado. 1 Ha fracasado más de lo que debería. 2 Cuando mira atrás, mira fracaso tras fracaso. 3 Se siente una persona totalmente fracasado.</p> <p>4. Pérdida de placer</p> <p>0 Disfruta de las cosas que le gustan tanto como antes. 1 No disfruta de las cosas tanto como antes. 2 Obtiene muy poco placer de las cosas con las que antes disfrutaba. 3 No obtiene ningún placer de las cosas con las que antes disfrutaba.</p> <p>5. Sentimientos de culpa</p> <p>0 No se siente especialmente culpable. 1 Se siente culpable de muchas cosas que ha hecho o debería haber hecho. 2 Se siente bastante culpable la mayor parte del tiempo. 3 Se siente culpable consistentemente.</p>	<p>6. Sentimientos de castigo</p> <p>0 No siente que está siendo castigado. 1 Siente que puede ser castigado. 2 Espera ser castigado. 3 Siente que está siendo castigado.</p> <p>7. Inestabilidad con uno mismo</p> <p>0 Siente lo mismo que antes sobre usted mismo. 1 Ha perdido confianza en sí mismo. 2 Está desconfiado consigo mismo. 3 No se gusta.</p> <p>8. Autoexigencia</p> <p>0 No se critica o se culpa más que antes. 1 Es más crítico consigo mismo de lo que solía ser. 2 Critica todos sus defectos. 3 Se culpa por todo lo malo que sucede.</p> <p>9. Pensamientos o deseos de suicidio</p> <p>0 No tiene ningún pensamiento de suicidio. 1 Tiene pensamientos de suicidio, pero no los llevaría a cabo. 2 Le gustaría suicidarse. 3 Se suicidaría si tuviese la oportunidad.</p> <p>10. Llanto</p> <p>0 No llora más de lo que solía hacerlo. 1 Lloro más de lo que solía hacerlo. 2 Lloro por cualquier cosa. 3 Tiene ganas de llorar continuamente, pero no puede.</p>
<p>1. Tristeza</p> <p>0 No se siente triste habitualmente. 1 Se siente triste gran parte del tiempo. 2 Se siente triste continuamente. 3 Se siente tan triste o tan desgraciado que no puede soportarlo.</p> <p>2. Pesimismo</p> <p>0 No está desanimado sobre su futuro. 1 Se siente más desanimado sobre su futuro que antes. 2 No espera que las cosas mejoren. 3 Siente que su futuro es desesperanzador y que las cosas sólo empeorarán.</p> <p>3. Sentimientos de fracaso</p> <p>0 No se siente fracasado. 1 Ha fracasado más de lo que debería. 2 Cuando mira atrás, mira fracaso tras fracaso. 3 Se siente una persona totalmente fracasado.</p> <p>4. Pérdida de placer</p> <p>0 Disfruta de las cosas que le gustan tanto como antes. 1 No disfruta de las cosas tanto como antes. 2 Obtiene muy poco placer de las cosas con las que antes disfrutaba. 3 No obtiene ningún placer de las cosas con las que antes disfrutaba.</p> <p>5. Sentimientos de culpa</p> <p>0 No se siente especialmente culpable. 1 Se siente culpable de muchas cosas que ha hecho o debería haber hecho. 2 Se siente bastante culpable la mayor parte del tiempo. 3 Se siente culpable consistentemente.</p>	<p>6. Sentimientos de castigo</p> <p>0 No siente que está siendo castigado. 1 Siente que puede ser castigado. 2 Espera ser castigado. 3 Siente que está siendo castigado.</p> <p>7. Inestabilidad con uno mismo</p> <p>0 Siente lo mismo que antes sobre usted mismo. 1 Ha perdido confianza en sí mismo. 2 Está desconfiado consigo mismo. 3 No se gusta.</p> <p>8. Autoexigencia</p> <p>0 No se critica o se culpa más que antes. 1 Es más crítico consigo mismo de lo que solía ser. 2 Critica todos sus defectos. 3 Se culpa por todo lo malo que sucede.</p> <p>9. Pensamientos o deseos de suicidio</p> <p>0 No tiene ningún pensamiento de suicidio. 1 Tiene pensamientos de suicidio, pero no los llevaría a cabo. 2 Le gustaría suicidarse. 3 Se suicidaría si tuviese la oportunidad.</p> <p>10. Llanto</p> <p>0 No llora más de lo que solía hacerlo. 1 Lloro más de lo que solía hacerlo. 2 Lloro por cualquier cosa. 3 Tiene ganas de llorar continuamente, pero no puede.</p>		

Puntuación página 1 _____

Continúa en la página siguiente

BDI / 5


6 / Evaluación psicológica

Escala de Autoevaluación de Depresión (EAMD)-Zung

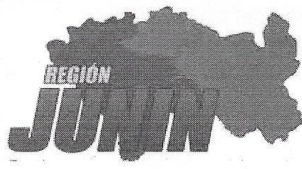
Escala Zung de Depresión (EAMD)	AyN:	Muy Pocas Veces	Algunas Veces	Muchas Veces	Casi Siempre
	Edad: G.I: Sexo:				
	Estado Civil:				
1.- Se siente triste y decaído.		1	2	3	4
2.- Por las mañanas se siente mejor.		4	3	2	1
3.- Tiene ganas de llorar y a veces llora.		1	2	3	4
4.- Le cuesta mucho dormir por la noche.		1	2	3	4
5.- Come igual que antes.		4	3	2	1
6.- Aún tiene deseos sexuales.		4	3	2	1
7.- Nota que está adelgazando.		1	2	3	4
8.- Está estropeado.		1	2	3	4
9.- El corazón le late más rápido que antes.		1	2	3	4
10.- Se cansa sin motivo.		4	3	2	1
11.- Su mente está tan despejada como siempre.		4	3	2	1
12.- Hace las cosas con la misma facilidad que antes.		4	3	2	1
13.- Se siente intranquilo y no puede mantenerse quieto.		1	2	3	4
14.- Tiene confianza en el futuro.		4	3	2	1
15.- Está más irritable que antes.		1	2	3	4
16.- Encuentra fácil tomar decisiones.		4	3	2	1
17.- Siente que es útil y necesario.		4	3	2	1
18.- Encuentra agradable vivir.		4	3	2	1
19.- Cree que sería mejor para los demás si estuviera muerto.		1	2	3	4
20.- Le gustan las mismas cosas que antes.		4	3	2	1
Sumatoria=					

Escalas Zung de Depresión (EAM)
Adecuación: CADE

Anexo 4: Constancia de Aplicación



GOBIERNO REGIONAL JUNÍN
HOSPITAL R.D.C.Q. "DANIEL A. CARRIÓN" - HYO
OFICINA DE APOYO A LA DOCENCIA E INVESTIGACIÓN



"Año de la lucha contra la Corrupción y la Impunidad"

Huancayo, 16 de abril del 2019.

CARTA N° 105 -2019-HRDCQ-DAC-HYO-OACDEL

Señor:
Dr. Carlos Alfredo, MENDOZA AMAYA.
JEFE DEL DEPARTAMENTO DE SALUD MENTAL.

PRESENTE.-


ASUNTO: TRABAJO DE INVESTIGACION.

REFERENTE: FORMULARIO UNICO DE TRAMITE CON EXPEDIENTE N°2193095.


Por medio de la presente es grato dirigirme a Ud., para saludarlo cordialmente, a nombre del Hospital Regional Docente Clínico Quirúrgico Daniel A., Carrión y la Oficina de Apoyo a la Capacitación Docencia e Investigación, en atención al documento de la referencia; con la **Evaluación y Aprobación del Presidente del Comité de Investigación del Hospital** y el visto bueno de la jefatura de la Oficina de Apoyo a la Capacitación Docencia e Investigación, comunico la **AUTORIZACION**, del uso de campo clínico, para la **ejecución**, de tesis titulado **"ESTUDIO DE PRECISION DIAGNOSTICA PARA EL INVENTARIO DE DETECCION CLINICA DE BECK (BDI-II), EN ADULTOS-2019"**, para optar título Profesional de **PSICOLOGO**, del Bachiller don: **Henry Antony, CADENAS HUIZA**, alumno de la Universidad Peruana los Andes.

En espera de la atención a la presente, solicito brindarle las facilidades del caso, reciba Ud., las muestras de mi especial consideración y estima personal.

Atentamente.



M. DE LA CRUZ



DIRECCIÓN GENERAL

GOBIERNO REGIONAL JUNÍN
HOSPITAL REGIONAL DOCENTE CLINICO
QUIRURGICO DANIELA CARRION - HUANCAYO

[Signature]

Dr. Juan Carlos Vergara Quintanilla
DIRECTOR ADJUNTO
CMP: 26852

Cc
Archivo.
MGVL/PGR
MDS/nmmr
Interesado(a)

HRDCQ "DAC" - HYO	
REG N°	3266439
EXP. N°	2173095

AV. DANIEL A. CARRIÓN N° 1550-1552-TELEFAX 064-222157

Anexo 5: Confiabilidad del Instrumento

El Inventario de Depresión de Beck, es un instrumento clínico de detección, compuesto por 21 reactivos orientados a la exploración sintomatológica de depresión, el instrumento está compuesto por 2 dimensiones, el primero referido a síntomas Cognitivos (el cual agrupa a los siguientes reactivos: 1,2,3,4,5,6,7,8,9,12,13,20 y 21) con indicadores como: tristeza, indecisión, sentimiento de culpa e impulsos suicidas, la segunda dimensión confiere síntomas Fisiológicos (el cual agrupa a los siguientes reactivos: 10,11,14,15,16,17,18 y 19) con indicadores como: periodos de llanto, imagen corporal, pérdida de apetito, pérdida de peso e irritabilidad. En el análisis de su confiabilidad obtuvo un coeficiente de 0,79 para la escala global, lo cual implica que demuestra aceptable homogeneidad entre todos los reactivos que la componen, por otro lado, obtuvo un coeficiente de 0,69 para dimensión Cognitiva y 0,57 para la dimensión Fisiológica, a continuación, se condensan los resultados en la tabla 41 y en la tabla 42 se muestra la consistencia interna por cada reactivo.

Tabla 41
Alfa de Cronbach BDI-II (escala completa y dimensiones).

Alfa de Cronbach		Cociente de Fiabilidad	Número de Elementos Analizados	Grado de Homogeneidad/Criterios de George y Mallery (2003)
Análisis Central	<i>Escala Global</i>	0,79	21	<i>Aceptable</i>
Dimensiones	<i>Cognitiva</i>	0,69	13	<i>Cuestionable</i>
	<i>Fisiológica</i>	0,57	8	<i>Pobre</i>

Fuente: Cadenas, H. (2019). Recurso Personal (Cálculo SPSS).

Tabla 42
Análisis de Fiabilidad BDI-II (por reactivos).

	Media de escala si el elemento se ha suprimido	Varianza de escala si el elemento se ha suprimido	Correlación total de elementos corregida	Alfa de Cronbach si el elemento se ha suprimido
Llanto	14,86	55,184	,661	,763
Agitación	14,90	62,867	,215	,792
Inutilidad	15,38	64,077	,189	,792
Pérdida de energía	14,90	61,684	,276	,789

Cambios en el patrón de sueño	14,64	58,970	,373	,784
Irritabilidad	15,14	60,613	,439	,781
Cambios en el apetito	14,94	61,486	,305	,787
Dificultad de concentración	15,02	62,796	,274	,789
Tristeza	14,72	57,512	,476	,776
Pesimismo	15,10	62,133	,361	,785
Sentimiento de fracaso	15,04	60,570	,418	,781
Pérdida de placer	14,58	60,167	,372	,783
Sentimiento de culpa	14,90	60,704	,360	,784
Sentimiento de castigo	15,18	60,396	,289	,790
Insatisfacción con uno mismo	14,90	57,847	,519	,774
Autocríticas	14,84	61,566	,257	,791
Pensamientos o deseos de suicidio	15,38	62,771	,304	,788
Pérdida de interés	15,02	62,714	,267	,789
Indecisión	14,88	58,679	,416	,781
Cansancio o fatiga	14,98	61,326	,364	,784
Pérdida de interés en el sexo	15,10	63,194	,165	,796

Fuente: Cadenas, H. (2019). Recurso Personal (Cálculo SPSS).

Además de ser un instrumento fiable, el Inventario de Depresión de Beck también demuestra buen grado de reproductibilidad, por lo que para la concordancia inter-examinadores el BDI-II demostró un coeficiente de replicación equivalente al 0,72 lo cual bajo los criterios de Tornimbeni, Pérez y Olaz propuestos el 2008, el BDI-II demostraría una fuerza de concordancia considerada como BUENA.

Anexo 6: Data del Procesamiento de Datos

La base de datos que se presenta a continuación integra todos los resultados tabulados durante el proceso de ejecución [210 casos/muestra apareada], inicialmente se integró los datos de Excel [Edad, Sexo, EC y GI] con los datos tabulados en SPSS [Estado Actual, EAMD, GSDicotómico, BDI-II y PEDicotómico], posteriormente todos los datos agrupados fueron exportados a un libro en Excel para poder presentarlas de esta manera, se especifican a continuación las variables para cada valor; Sexo [1: Varón/2: Mujer], Estado Civil [1: Soltero/2: Con pareja/3: Conviviente/4: Casado/5: Divorciado/6: Viudo], Instrucción [1: Primaria (C/I)/2: Secundaria (C/I)/3: Sup. Técnico (C/I)/4: Superior (C/I)/5: Iltrado], Estado Actual [0: Cond. Clínica Ausente/1: Cond. Clínica Presente], GSDicotómico [Gold Standard resultados agrupados; 0: Verdadero/1: Falso] y PEDicotómico [Prueba en Estudio resultados agrupados; 0: Negativo/1: Positivo].

Edad	Sexo	EstadoCivil	Instrucción	EA	EAMD	GSDicotómico	BDI-II	PEDicotómico
23	2	1	4	0	33	1	2	0
50	2	1	4	0	29	0	1	0
55	2	4	3	1	45	1	25	1
34	1	1	4	1	33	1	10	0
44	2	4	2	1	38	1	15	1
28	2	1	4	1	51	1	21	1
34	2	1	4	1	54	1	27	1
20	2	1	4	1	39	1	16	1
38	2	3	2	0	48	1	15	1
19	2	1	2	1	60	1	20	1
53	2	6	2	1	64	1	29	1
21	1	2	3	1	62	1	29	1
53	2	6	4	1	58	1	18	1
20	2	1	4	1	68	1	42	1
19	2	1	2	1	58	1	25	1
31	1	3	4	1	29	0	10	0
28	1	1	4	0	26	0	9	0
50	2	4	2	0	51	1	8	0
56	2	5	4	1	63	1	17	1
28	1	1	2	0	67	1	40	1
22	2	2	4	1	29	0	15	1
49	2	4	4	0	26	0	1	0
45	2	1	2	0	28	0	1	0

64	2	6	5	1	57	1	20	1
53	1	3	2	0	25	0	3	0
22	2	1	4	0	26	0	4	0
21	2	1	4	1	36	1	19	1
48	1	4	4	0	29	0	4	0
21	2	1	4	0	36	1	10	0
25	2	1	4	1	35	1	17	1
21	2	2	4	0	24	0	3	0
20	2	1	4	0	27	0	2	0
40	1	1	4	1	35	1	2	0
21	2	1	4	1	54	1	43	1
21	2	1	4	0	30	0	3	0
21	2	2	4	0	30	0	6	0
23	2	3	4	1	43	1	15	1
59	2	3	4	1	46	1	19	1
25	2	1	4	1	43	1	17	1
59	1	3	1	1	47	1	23	1
55	2	4	2	1	54	1	15	1
62	1	4	1	0	26	0	3	0
52	2	3	4	0	22	0	6	0
44	2	4	4	1	38	1	14	1
42	1	1	4	1	34	1	11	1
34	1	1	2	1	31	0	12	1
28	1	3	1	1	51	1	18	1
41	2	3	3	0	58	1	9	0
45	2	3	4	0	25	0	5	0
56	2	4	1	1	54	1	27	1
61	1	4	2	1	26	0	12	1
19	2	1	4	1	28	0	13	1
36	1	3	2	1	62	1	37	1
44	2	3	3	1	54	1	14	1
29	2	2	4	0	38	1	7	0
21	2	1	4	1	48	1	15	1
24	2	2	4	1	54	1	23	1
18	1	1	2	0	46	1	6	0
29	1	2	3	1	32	0	22	1
45	2	3	1	1	49	1	11	1
74	1	6	1	1	55	1	5	0
56	1	4	5	0	26	0	4	0
25	1	2	3	0	27	0	8	0
60	2	3	1	1	54	1	13	1
32	1	4	4	1	36	1	16	1
28	1	2	3	0	29	0	6	0
71	2	5	2	0	54	1	37	1
56	2	3	2	0	35	1	7	0
23	2	2	4	1	34	1	22	1

28	2	1	4	1	51	1	35	1
18	1	1	3	1	26	0	8	0
43	1	3	2	1	47	1	18	1
62	2	4	2	1	48	1	23	1
57	2	3	1	1	26	0	15	1
29	2	2	3	0	54	1	1	0
22	2	1	4	1	30	0	9	0
35	2	1	4	1	29	0	42	1
41	1	3	3	0	28	0	14	1
56	2	4	5	0	48	1	2	0
19	2	1	4	0	62	0	8	0
19	2	1	4	1	49	1	22	1
53	2	1	4	0	29	0	7	0
56	1	3	2	1	35	1	17	1
18	2	1	2	1	33	1	24	1
48	2	4	4	1	48	1	21	1
63	2	5	1	1	35	1	12	1
22	2	1	4	1	48	1	14	1
34	1	4	3	1	27	0	9	0
58	2	4	1	0	22	0	10	1
56	2	5	4	0	26	0	6	0
23	2	1	4	1	64	1	27	1
39	1	1	3	1	42	1	23	1
18	2	1	2	0	71	1	10	1
50	1	3	1	1	34	1	26	1
21	1	2	4	0	49	0	6	0
19	2	1	4	0	32	0	6	0
33	2	3	2	0	20	0	4	0
59	2	5	1	1	54	1	18	1
50	2	4	2	1	51	1	12	1
19	1	1	2	1	70	1	17	1
24	2	1	3	0	21	0	10	0
36	2	4	4	0	26	0	4	0
22	1	1	4	1	62	1	16	1
19	2	1	3	0	33	1	10	0
39	1	1	4	1	34	1	18	1
55	2	3	1	1	60	1	26	1
27	2	4	4	1	52	1	34	1
42	1	3	3	0	30	0	3	0
45	2	3	4	0	21	0	7	0
28	2	5	4	1	38	1	27	1
20	2	1	4	1	70	1	23	1
47	2	4	2	1	35	1	25	1
19	2	1	2	1	54	1	13	1
38	2	4	2	0	34	1	8	0
40	1	4	2	0	25	0	2	0

21	2	1	4	0	27	0	9	0
35	2	3	2	0	20	0	8	0
42	2	4	3	1	33	1	14	1
51	2	4	2	1	48	1	12	1
28	1	1	3	1	47	1	34	1
34	1	3	2	0	22	0	1	0
57	2	4	1	1	37	1	32	1
26	2	3	4	1	41	1	7	0
28	1	1	3	1	50	1	24	1
22	2	2	4	0	26	0	5	0
67	2	4	1	1	21	0	9	0
56	2	3	2	1	48	1	43	1
23	2	1	4	1	56	1	8	0
18	2	1	2	1	63	1	10	0
73	2	6	1	1	60	1	12	1
21	2	1	4	1	44	1	16	1
20	1	1	2	0	23	0	8	0
33	2	2	3	0	29	0	6	0
59	2	4	1	1	51	1	12	1
50	1	4	2	0	29	0	8	0
29	1	1	2	1	34	1	13	1
24	2	3	3	0	28	0	10	0
36	2	1	4	0	21	0	7	0
22	2	1	2	0	29	0	2	0
24	2	1	2	1	43	1	26	1
39	1	3	4	1	23	0	9	0
55	2	5	2	1	62	1	1	0
36	2	2	4	0	37	1	7	0
20	2	1	2	0	32	0	10	0
47	1	3	3	0	34	1	15	1
18	2	1	4	1	26	0	8	0
38	2	3	2	1	53	1	11	1
40	1	5	4	0	20	0	8	0
21	2	2	4	0	22	0	4	0
35	2	2	4	1	39	1	16	1
42	2	3	4	1	37	1	16	1
18	1	1	2	1	21	0	9	0
29	2	3	3	1	52	1	11	1
34	2	1	3	0	46	1	17	1
19	2	1	4	0	39	1	19	1
38	2	2	4	0	45	1	10	0
28	2	1	4	1	67	1	13	1
64	1	6	2	0	23	0	8	0
34	1	1	4	1	28	0	16	0
30	2	1	4	1	33	1	26	1
28	2	1	4	0	26	0	8	0

20	2	1	4	0	23	0	1	0
61	1	4	1	0	31	0	5	0
52	2	4	1	0	21	0	10	0
34	2	2	2	1	38	1	13	1
57	2	3	2	1	23	0	8	0
26	1	1	4	1	45	1	4	0
28	2	1	4	1	61	1	9	0
22	1	2	2	1	34	1	14	1
67	2	3	5	1	46	1	13	1
56	2	3	2	0	30	0	7	0
23	2	1	2	1	29	0	5	0
18	1	1	2	1	33	1	15	1
73	2	6	1	1	28	0	9	0
21	2	1	2	0	41	1	8	0
22	1	1	4	0	33	1	12	1
33	1	2	4	1	40	1	14	1
59	2	4	3	0	57	1	11	1
50	2	4	1	1	37	1	16	1
61	2	4	2	0	50	1	19	1
52	1	3	2	1	25	0	6	0
34	1	3	4	1	29	0	5	0
57	1	4	3	1	30	0	8	0
26	2	1	4	1	52	1	15	1
28	2	1	2	0	24	0	9	0
18	2	1	2	1	48	1	22	1
22	2	1	2	1	28	0	14	1
30	2	2	2	0	31	0	2	0
29	2	1	4	0	39	1	7	0
21	1	1	2	1	56	1	23	1
26	2	2	3	1	21	0	7	0
43	2	3	4	1	38	1	22	1
56	1	4	2	0	32	0	6	0
20	2	1	4	0	28	0	8	0
37	2	2	4	1	22	0	26	1
58	1	3	1	1	24	0	2	0
20	2	1	3	1	47	1	11	1
33	2	3	2	1	35	1	14	1
48	1	3	3	1	21	0	8	0
60	2	4	1	1	41	1	22	1
51	2	4	1	0	25	0	9	0
29	2	2	2	1	32	0	15	1
22	2	1	4	1	34	1	10	0
31	2	2	4	1	39	1	18	1
58	2	3	3	1	61	1	5	0
35	1	3	4	1	41	1	12	1
22	2	2	4	1	52	1	26	1

19	2	2	4	1	23	0	13	1
36	2	3	4	0	28	0	2	0
44	2	3	4	0	31	0	8	0

Anexo 7: Consentimiento Informado

CONSENTIMIENTO INFORMADO

El presente documento de carácter ético, académico e informativo mantiene la intencionalidad de incluir a su persona a la investigación clínica ***“Estudio de precisión diagnóstica para el inventario de detección clínica de Depresión de Beck (BDI-II) en adultos, Huancayo-2019”***, la cual se llevará a cabo en el presente año por medio del investigador: CADENAS HUIZA Henry A., egresado de carrera profesional de Psicología.

- Autorizando con el presente consentimiento su *“participación voluntaria y anónima”* en el presente estudio clínico, el cual tiene como objetivo: ***“Validar pruebas diagnósticas para Depresión a través del análisis de su capacidad discriminativa y ofrecer mayor seguridad respecto a la utilidad de las mismas en su proceso de evaluación”***.
- Recuerde que su participación es tan o más importante que la misma investigación, de allí su valor aplicativo, puesto que además de formar parte de un proceso de investigación clínica que generará avances en el conocimiento de la Psicología clínica y de la salud, primordialmente representará para usted a posteriori mejoras para el procedimiento de evaluación psicológica y monitorización clínica.

Dejando claro entonces y recalcando el carácter voluntario/anónimo de su participación, así como la posibilidad de retirarse del estudio en cualquier momento del proceso, sin que ello altere la relación Clínico-Paciente, ni mucho menos se produzca perjuicio para con su atención en el presente establecimiento de salud.

Yo: por medio de la presente accedo a participar de manera voluntaria y anónima en la presente investigación.

- He podido hacer preguntas sobre el estudio.
- He recibido suficiente información sobre el estudio.

Comprendo, por tanto, que mi participación es voluntaria, por ende:

1. Se mantendrá confidencialidad de mi información personal (edad, sexo, etc.) no autorizada a la divulgación pública de las mismas.

Comprendo que puedo retirarme del estudio:

- Cuando quiera.
- Sin tener que dar explicaciones.
- Sin que esto repercuta en mis cuidados médicos.
- Sin que esto afecte mi bienestar integral.

Huancayo, _____ del 2018.

.....
Firma de conformidad
del participante.

.....
Firma de acuerdo del
investigador.

PUBLICACIÓN DE RESULTADOS CLÍNICOS

El presente documento de carácter ético, académico e informativo tiene como objetivo educar a usted acerca de la importancia que conlleva la publicación de resultados producto de procesos de investigación, puesto que gracias ello la comunidad académica juzgará la utilidad de los mismos y la población en general se beneficiará de sus resultados en términos de mejoras para con la intervención que realizan los profesionales con su salud.

Investigación Clínica: *“Estudio de precisión diagnóstica para el inventario de detección clínica de Depresión de Beck (BDI-II) en adultos, Huancayo-2019”*.

Yo:he sido informado(a) sobre los especialistas que tendrán acceso a los datos expedidos voluntariamente por mi persona.

- He podido hacer acerca de la publicación de resultados.
- He recibido información pertinente sobre el carácter de dicho estudio.

Comprendo que mis resultados:

1. Serán expuestas de manera anónima.
2. Salvaguardando mi bienestar propio.
3. Serán expuestos a una comunidad científica.
4. Serán analizados científicamente.
5. Mantendrán la confidencialidad en el tiempo.

Comprendo nuevamente entonces que puedo retirarme del estudio:

- Cuando quiera.
- Sin tener que dar explicaciones.
- Sin que esto repercuta en mis cuidados médicos.
- Sin que esto afecte mi bienestar integral.

Por tanto, presto voluntariamente mi conformidad para la presentación del análisis de mis resultados anónimos no siendo presentados como caso único sino aunados al resto de pacientes como participantes.

Huancayo, _____ del 2018.

.....
Firma de conformidad
del participante.

.....
Firma de acuerdo del
investigador.

Anexo 8: Evidencias de la Aplicación del Instrumento

BDE-II

Nombre: Edgar S. Sosa Estado civil: Casado Profesión: Psicólogo

Instrucciones: Este cuestionario consiste en 21 grupos de afirmaciones. Por favor, lee con atención cada uno de ellos y, a continuación, señala cuál de las afirmaciones de cada grupo describe mejor el modo en el que se ha sentido DURANTE LAS DOS ÚLTIMAS SEMANAS, INCLUYENDO EL DÍA DE HOY. Rodea con un círculo el número que se encuentre escrito a la izquierda de la afirmación que haya elegido. Si dentro del mismo grupo hay más de una afirmación que considere igualmente aplicable a su caso, señalada también. Asegúrese de leer todas las afirmaciones dentro de cada grupo antes de efectuar la elección.

<p>1. Tristeza</p> <p>0 No se siente triste habitualmente.</p> <p>1 Se siente triste gran parte del tiempo.</p> <p>2 Se siente triste continuamente.</p> <p>3 Se siente tan triste o tan deprimido que no puede soportarlo.</p> <p>2. Peleleísmo</p> <p>0 No está demasiado sobre su futuro.</p> <p>1 Se preocupa mucho por el futuro que antes.</p> <p>2 No espera que las cosas mejoren y que las cosas solo empeoren.</p> <p>3 Siéntese que su futuro es desalentador y que las cosas solo empeorarán.</p> <p>3. Sentimientos de fracaso</p> <p>0 No se siente fracasado.</p> <p>1 Ha fracasado más de lo que debería.</p> <p>2 Cuando mira atrás, siente fracaso y frustración.</p> <p>3 Se siente una persona totalmente fracasada.</p> <p>4. Pérdida de placer</p> <p>0 Disfruta de las cosas que le gustan tanto como antes.</p> <p>1 No disfruta de las cosas que antes disfrutaba.</p> <p>2 Obtiene muy poco placer de las cosas con las que antes disfrutaba.</p> <p>3 No obtiene ningún placer de las cosas con las que antes disfrutaba.</p> <p>5. Sentimientos de culpa</p> <p>0 No se siente especialmente culpable.</p> <p>1 Se siente culpable de muchas cosas que ha hecho o debería haber hecho.</p> <p>2 Se siente bastante culpable la mayor parte del tiempo.</p> <p>3 Se siente culpable constantemente.</p>	<p>6. Sentimientos de castigo</p> <p>0 No siente que está siendo castigado.</p> <p>1 Siente que puede ser castigado.</p> <p>2 Espera ser castigado.</p> <p>3 Siente que está siendo castigado.</p> <p>7. Insatisfacción con uno mismo</p> <p>0 Siente lo mismo que antes sobre usted mismo.</p> <p>1 Ha perdido confianza en sí mismo.</p> <p>2 Está descontento consigo mismo.</p> <p>3 No se gusta.</p> <p>8. Autocrítica</p> <p>0 No se critica o se culpa más que antes.</p> <p>1 Es más crítico con uno mismo de lo que solía ser.</p> <p>2 Critica todos sus defectos.</p> <p>3 Se culpa por todo lo malo que sucede.</p> <p>9. Pensamientos o deseos de suicidio</p> <p>0 No tiene ningún pensamiento de suicidio.</p> <p>1 Tiene pensamientos de suicidio, pero no los llevaría a cabo.</p> <p>2 Se gustaría suicidarse.</p> <p>3 Se suicidaría si tuviera la oportunidad.</p> <p>10. Llanto</p> <p>0 No llora más de lo que solía llorar.</p> <p>1 Lloro más de lo que solía llorar.</p> <p>2 Lloro por cualquier cosa.</p> <p>3 Tiene ganas de llorar continuamente, pero no puede.</p>
---	--

17. Irritabilidad

0 No está más irritable de lo habitual.

1 Está más irritable de lo habitual.

2 Está mucho más irritable de lo habitual.

3 Está irritable continuamente.

18. Cambios en el apetito

0 No ha experimentado ningún cambio en su apetito.

1 Su apetito es algo menor de lo habitual.

2 Su apetito es algo mayor de lo habitual.

3 Su apetito es mucho mayor que antes.

4 Su apetito es mucho mayor de lo habitual.

5 Ha perdido completamente el apetito.

6 Tiene ganas de comer continuamente.

19. Dificultad de concentración

0 Puede concentrarse tan bien como siempre.

1 No puede concentrarse tan bien como habitualmente.

2 Le cuesta mantenerse concentrado en algo durante mucho tiempo.

3 No puede concentrarse en nada.

20. Camarero o fatiga

0 No está más cansado o fatigado que de costumbre.

1 Se cansa o fatiga más fácilmente que de costumbre.

2 Está demasiado cansado o fatigado para hacer muchas cosas.

3 Está demasiado cansado o fatigado para hacer la mayoría de las cosas que antes solía hacer.

21. Pérdida de interés en el sexo

0 No ha notado ningún cambio reciente en su interés por el sexo.

1 Está menos interesado por el sexo de lo que solía estar.

2 Está mucho menos interesado por el sexo ahora.

3 Ha perdido completamente el interés por el sexo.

11. Agresión

0 No está más inquieto o agitado que de costumbre.

1 Está inquieto o agitado que de costumbre.

2 Está tan inquieto o agitado que tiene que estar continuamente moviéndose o haciendo algo.

3 Está tan inquieto o agitado que tiene que estar continuamente moviéndose o haciendo algo.

12. Pérdida de interés

0 No ha perdido el interés por otras personas o actividades.

1 Está menos interesado que antes por otras personas o actividades.

2 Está mucho menos interesado que antes por otras personas o actividades.

3 Le resulta difícil interesarse en algo.

13. Indecisión

0 Toma decisiones más o menos como siempre.

1 Tomar decisiones le resulta más difícil que de costumbre.

2 Tiene mucha más dificultad en tomar decisiones que de costumbre.

3 Tiene problemas para tomar cualquier decisión.

14. Inutilidad

0 No se siente inútil.

1 No se considera tan valioso y útil como solía ser.

2 Se siente inútil en comparación con otras personas.

3 Se siente completamente inútil.

15. Pérdida de energía

0 Tiene tanta energía como siempre.

1 Tiene menos energía de la que solía tener.

2 No tiene suficiente energía para hacer muchas cosas.

3 No tiene suficiente energía para hacer nada.

16. Cambios en el patrón de sueño

0 No ha experimentado ningún cambio en su patrón de sueño.

1 a Duermo algo más de lo habitual.

1 b Duermo algo menos de lo habitual.

2 a Duermo mucho más de lo habitual.

2 b Duermo mucho menos de lo habitual.

3 a Se despierta a mayor parte del día.

3 b Se despierta a mayor parte de la noche y no puede volver a dormir.

1.4 Puntuación página 1

4.4 Puntuación página 2

5.8 Puntuación total (págs. 1 + 2)

6 / Evaluación psicológica

BDI / 5

BDI-II

Nombre: ... Estado civil: ... Profesión: ... Edad: ... Sexo: ...

Instrucciones: Este cuestionario consiste en 21 grupos de afirmaciones. Por favor, lee con atención cada uno de ellos y a continuación, señala cuál de las afirmaciones de cada grupo describe mejor el modo en el que se ha sentido durante las DOS ÚLTIMAS SEMANAS, INCLUYENDO EL DÍA DE HOY. Rodée con un círculo el número de la afirmación que mejor describe la intensidad de la afirmación que hayas elegido. Si dentro del mismo grupo hay más de una afirmación que consideras igualmente adecuada, puedes rodearlas con un círculo. Anota el número de las afirmaciones dentro de cada grupo antes de efectuar la elección.

- 1. Tristeza
2. Pérdida de interés
3. Sentimientos de culpa
4. Sentimientos de desesperanza
5. Sentimientos de ira
6. Sentimientos de castigo
7. Insatisfacción con uno mismo
8. Autoresistencia
9. Pensamientos o deseos de suicidio
10. Llanto
11. Tristeza
12. Pérdida de interés
13. Sentimientos de culpa
14. Sentimientos de desesperanza
15. Sentimientos de ira
16. Sentimientos de castigo
17. Insatisfacción con uno mismo
18. Autoresistencia
19. Pensamientos o deseos de suicidio
20. Llanto

Puntuación página 1

Continúa en la página siguiente

BDI / 5

- 11. Agitación
12. Pérdida de interés
13. Indecisión
14. Inutilidad
15. Pérdida de energía
16. Cambios en el patrón de sueño
17. Irritabilidad
18. Cambios en el apetito
19. Dificultad de concentración
20. Cambiado o fallido
21. Pérdida de interés en el sexo

Puntuación página 1

Puntuación página 2

Puntuación total (págs. 1 + 2)

6 / Evaluación psicológica

AyN: <u>Ever Córtez Atachekua</u>		Muy Pocas Veces	Algunas Veces	Muchas Veces	Casi Siempre
Edad: <u>55a</u> G.I: <u>Sup</u> Iq. Sexo: <u>♀</u>					
Estado Civil: <u>Casada</u>					
Escala Zung de Depresión (EAMD)	1.- Se siente triste y decaído.	1	2 <input checked="" type="checkbox"/>	3	4
	2.- Por las mañanas se siente mejor.	4	3 <input checked="" type="checkbox"/>	2	1
	3.- Tiene ganas de llorar y a veces llora.	1	2 <input checked="" type="checkbox"/>	3	4
	4.- Le cuesta mucho dormir por la noche.	1	2	3 <input checked="" type="checkbox"/>	4
	5.- Come igual que antes.	4	3 <input checked="" type="checkbox"/>	2	1
	6.- Aún tiene deseos sexuales.	4 <input checked="" type="checkbox"/>	3	2	1
	7.- Nota que está adelgazando.	1	2 <input checked="" type="checkbox"/>	3	4
	8.- Está estreñido.	1 <input checked="" type="checkbox"/>	2	3	4
	9.- El corazón le late más rápido que antes.	1 <input checked="" type="checkbox"/>	2	3	4
	10.- Se cansa sin motivo.	4	2	3	4 <input checked="" type="checkbox"/>
	11.- Su mente está tan despejada como siempre.	4	3	2 <input checked="" type="checkbox"/>	1
	12.- Hace las cosas con la misma facilidad que antes.	4	3 <input checked="" type="checkbox"/>	2	1
	13.- Se siente intranquilo y no puede mantenerse quieto.	1	2	3 <input checked="" type="checkbox"/>	4
	14.- Tiene confianza en el futuro.	4	3	2	1 <input checked="" type="checkbox"/>
	15.- Está más irritable que antes.	1	2	3 <input checked="" type="checkbox"/>	4
	16.- Encuentra fácil tomar decisiones.	4	3	2 <input checked="" type="checkbox"/>	1
	17.- Siente que es útil y necesario.	4	3	2	1
	18.- Encuentra agradable vivir.	4	3	2	1 <input checked="" type="checkbox"/>
	19.- Cree que sería mejor para los demás si estuviera muerto.	1 <input checked="" type="checkbox"/>	2	3	4
	20.- Le gustan las mismas cosas que antes.	4 <input checked="" type="checkbox"/>	3	2	1
Sumatoria=		11	15	13	06

$$\Sigma = 45 = 56$$

Escalas Zung de Depresión (EAMI)
Adecuación: CADEI

AyN: Jorge Javier Cordero Pa Edad: 34 G.I: Sup C Sexo: ♂ Estado Civil: Soltero		Muy Pocas Veces	Algunas Veces	Muchas Veces	Casi Siempre
		1.- Se siente triste y decaído.	1	2 X	3
2.- Por las mañanas se siente mejor.	4	3	2 X	1	
3.- Tiene ganas de llorar y a veces llora.	1	2 X	3	4	
4.- Le cuesta mucho dormir por la noche.	1 X	2	3	4	
5.- Come igual que antes.	4	3	2	1 X	
6.- Aún tiene deseos sexuales.	4	3	2	1 X	
7.- Nota que está adelgazando.	1 X	2	3	4	
8.- Está estreñido.	1 X	2	3	4	
9.- El corazón le late más rápido que antes.	1 X	2	3	4	
10.- Se cansa sin motivo.	4 X	3	2	1	
11.- Su mente está tan despejada como siempre.	4	3 X	2	1	
12.- Hace las cosas con la misma facilidad que antes.	4	3	2 X	1	
13.- Se siente intranquilo y no puede mantenerse quieto.	1 X	2	3	4	
14.- Tiene confianza en el futuro.	4	3	2	1 X	
15.- Está más irritable que antes.	1 X	2	3	4	
16.- Encuentra fácil tomar decisiones.	4	3	2 X	1	
17.- Siente que es útil y necesario.	4	3	2 X	1	
18.- Encuentra agradable vivir.	4	3 X	2	1	
19.- Cree que sería mejor para los demás si estuviera muerto.	1	2	3	4 X	
20.- Le gustan las mismas cosas que antes.	4	3	2	1 X	
Sumatoria=	07	10	08	08	

$$\Sigma = 33 = \text{Índice} = 41$$

Escalas Zung de Depresión (EAMD)
Adecuación: CADEN

HOSPITAL EDCO "DANIEL ALCIDES CARRIÓN"
Servicio de Psicología y Psiquiatría

CONSENTIMIENTO INFORMADO

El presente documento de carácter ético, académico e informativo manifiesta la idoneidad de incluir a su persona a la investigación clínica "Estudio de prescripción diagnóstica para el diagnóstico de depresión clínica de Beck (BDI-II) en adultos, Huancayo-2019", la cual se llevará a cabo en el presente año por medio del investigador: CADENAS HOLZA Henry A., egresado de carrera profesional de Psicología.

- Autorizo con el presente consentimiento su "participación voluntaria y autónoma" en el presente estudio clínico, el cual tiene como objetivos: "Validar protocolos diagnósticos para Depresión a través del análisis de su especificidad diagnóstica y ofrecer mayor seguridad respecto a la utilidad de los mismos en su proceso de evaluación".
- Reconozco que su participación es tan o más importante que la misma investigación, de allí su valor educativa, puesto que además de formar parte de un proceso de investigación clínica que generará avances en el conocimiento de la Psicología clínica y de la salud, proporcionaré soporte representativo para usar a posteriori mejoras para el procedimiento de evaluación psicológica y monitorización clínica.

Dejando claro entonces y reconociendo el carácter voluntario/autónomo de su participación, así como la posibilidad de retirarse del estudio en cualquier momento del proceso, sin que ello altere la relación Clínico-Paciente, ni mucho menos se produzca perjuicio para con su atención en el presente establecimiento de salud.

Yo, Sra. Carolina Ancochea, por medio de la presente suceso a participar de manera voluntaria y autónoma en la presente investigación.

- He podido hacer preguntas sobre el estudio.
- He recibido suficiente información sobre el estudio.

Comprendo, por tanto, que mi participación es voluntaria, por ende:

1. Se mantendrá confidencialidad de mi información personal (edad, sexo, etc.) no autorizada a la divulgación pública de las mismas.

Comprendo que puedo retirarme del estudio:

- Cuando quiera.
- Sin tener que dar explicaciones.
- Sin que esto repercuta en mis cuidados médicos.
- Sin que esto afecte mi bienestar integral.


Firma de conformidad del participante.

Huancayo 23 febrero del 2019.

Firma de acuerdo del investigador.

HOSPITAL EDCO "DANIEL ALCIDES CARRIÓN"
Servicio de Psicología y Psiquiatría

PUBLICACIÓN DE RESULTADOS CLÍNICOS

El presente documento de carácter ético, académico e informativo tiene como objetivo exponer a usted acerca de la importancia que conlleva la publicación de resultados producidos por procesos de investigación, puesto que gracias a la comunidad académica juzgará la utilidad de los mismos y la población en general se beneficiará de sus resultados en términos de mejoras para con la intervención que realicen los profesionales con su salud.

Investigación Clínica: "Estudio de prescripción diagnóstica para el diagnóstico clínico de Depresión de Beck (BDI-II) en adultos, Huancayo-2019".

Yo, Sra. Carolina Ancochea, he sido informada sobre los especialistas que tendrán acceso a los datos expeditos voluntariamente por mi persona.

- He podido hacer preguntas de la publicación de resultados.
- He recibido información pertinente sobre el carácter de dicho estudio.

Comprendo que mis realizaciones:

1. Serán expuestas de manera anónima.
2. Salvaguardando mi bienestar propio.
3. Serán expuestas a una comunidad científica.
4. Serán realizadas científicamente.
5. Mantendrán la confidencialidad en el tiempo.

Comprendo nuevamente entonces que puedo retirarme del estudio:

- Cuando quiera.
- Sin tener que dar explicaciones.
- Sin que esto repercuta en mis cuidados médicos.
- Sin que esto afecte mi bienestar integral.

Por tanto, prescindiendo voluntariamente de mi conformidad para la presentación del análisis de mis resultados anónimos no siendo presentados como tales datos sino suaves al resto de participantes ocupantes.


Firma de conformidad del participante.

Huancayo 23 febrero del 2019.

Firma de acuerdo del investigador.

HOSPITAL RDCQ "DANIEL ALCIDES CARRIÓN"
Servicio de Psicología y Psiquiatría

CONSENTIMIENTO INFORMADO

El presente documento de carácter ético, académico e informativo mantiene la intencionalidad de incluir a su persona a la investigación clínica "Estudio de precisión diagnóstica para el inventario de detección clínica de Depresión de Beck (BDI-II) en adultos, Huancayo-2019", la cual se llevará a cabo en el presente año por medio del investigador: CADENAS HUIZA Henry A., egresado de carrera profesional de Psicología.

- Autorizando con el presente consentimiento su "participación voluntaria y anónima" en el presente estudio clínico, el cual tiene como objetivo: "Validar pruebas diagnósticas para Depresión a través del análisis de su capacidad discriminativa y ofrecer mayor seguridad respecto a la utilidad de las mismas en su proceso de evaluación".
- Recuerde que su participación es tan o más importante que la misma investigación, de allí su valor aplicativo, puesto que además de formar parte de un proceso de investigación clínica que generará avances en el conocimiento de la Psicología clínica y de la salud, primordialmente representará para usted a posteriori mejoras para el procedimiento de evaluación psicológica y monitorización clínica.

Dejando claro entonces y recalcando el carácter voluntario/anónimo de su participación, así como la posibilidad de retirarse del estudio en cualquier momento del proceso, sin que ello altere la relación Clínico-Paciente, ni mucho menos se produzca perjuicio para con su atención en el presente establecimiento de salud.

Yo: Jorge Felipe Santos..... por medio de la presente accedo a participar de manera voluntaria y anónima en la presente investigación.

- He podido hacer preguntas sobre el estudio.
- He recibido suficiente información sobre el estudio.

Comprendo, por tanto, que mi participación es voluntaria, por ende:

1. Se mantendrá confidencialidad de mi información personal (edad, sexo, etc.) no autorizada a la divulgación pública de las mismas.

Comprendo que puedo retirarme del estudio:

- Cuando quiera.
- Sin tener que dar explicaciones.
- Sin que esto repercuta en mis cuidados médicos.
- Sin que esto afecte mi bienestar integral.

Huancayo, 13 marzo del 2019.


Firma de conformidad del participante.


Firma de acuerdo del investigador.

HOSPITAL RDCQ "DANIEL ALCIDES CARRIÓN"
Servicio de Psicología y Psiquiatría

PUBLICACIÓN DE RESULTADOS CLÍNICOS

El presente documento de carácter ético, académico e informativo tiene como objetivo educar a usted acerca de la importancia que conlleva la publicación de resultados producto de procesos de investigación, puesto que gracias ello la comunidad académica juzgará la utilidad de los mismos y la población en general se beneficiará de sus resultados en términos de mejoras para con la intervención que realizan los profesionales con su salud.

Investigación Clínica: "Estudio de precisión diagnóstica para el inventario de detección clínica de Depresión de Beck (BDI-II) en adultos, Huancayo-2019".

Yo: Jorge Felipe Santos..... he sido informado(s) sobre los especialistas que tendrán acceso a los datos expeditos voluntariamente por mi persona.

- He podido hacer acerca de la publicación de resultados.
- He recibido información pertinente sobre el carácter de dicho estudio.

Comprendo que mis resultados:

1. Serán expuestas de manera anónima.
2. Salvaguardando mi bienestar propio.
3. Serán expuestas a una comunidad científica.
4. Serán analizadas científicamente.
5. Mantendrán la confidencialidad en el tiempo.

Comprendo nuevamente entonces que puedo retirarme del estudio:

- Cuando quiera.
- Sin tener que dar explicaciones.
- Sin que esto repercuta en mis cuidados médicos.
- Sin que esto afecte mi bienestar integral.

Por tanto, presto voluntariamente mi conformidad para la presentación del análisis de mis resultados anónimos no siendo presentados como caso único sino aunados al resto de pacientes como participantes.

Huancayo, 13 marzo del 2019.


Firma de conformidad del participante.


Firma de acuerdo del investigador.