



Facultad de Ciencias de la Salud
Programa de Doctorado de Psicología

**Evaluación de resultados en rehabilitación
de personas con enfermedad mental grave:
Baremación del HoNOS y puntos de corte
para la toma de decisiones clínicas**

Tesis doctoral presentada por Nerea Iglesias López

Dirigida por el Dr. Ioseba Iraurgi Castillo

Firma doctoranda

Firma director

Nerea Iglesias López

Ioseba Iraurgi castillo

Bilbao, septiembre 2024

AGRADECIMIENTOS

En primer lugar, me gustaría darle las gracias a Ioseba, mi director de tesis, por acompañarme en los años de máster, prácticas y tesis. Gracias por ser mentor todos estos años y haberme dado la libertad y la confianza de desarrollarme en investigación, docencia y clínica, tres pilares que me apasionan. Gracias por guiarme y por los permisos que he necesitado estos años para poder iniciarme en esos tres pilares.

Gracias a Xabi, mi compañero de vida, por apoyarme sin dudar que lo conseguiría. Te has llevado la peor parte por los días malos que he tenido estos años, gracias por haber estado en cada uno de ellos a mi lado con cariño y comprensión.

Gracias aita y ama, por ser impulso, apoyo y fortaleza todos estos años desde que decidí ser psicóloga y todo lo que eso ha conllevado. Este trabajo no hubiera sido posible sin vosotros. A mi hermana Olaia, por todo el amor y apoyo que me das siempre, siempre. Gracias.

Gracias a mi familia por animarme y darme tanto cariño estos años, desde el principio hasta el final. Gracias a mis suegros Paco y Mari Mar por la confianza en el proceso y por tantos “tranquila, tu trabaja en la tesis lo que necesites, que nosotros nos encargamos de todo”.

Gracias a las Helenas (madre e hija), por acogerme y quererme tanto desde que nos conocimos y mantener el cariño aún en la distancia. Gracias por el apoyo en este proyecto, y en todos.

A mis amigas de Lapan, gracias por ser fortaleza y comprensión en las idas y venidas que he tenido en este proyecto, en todas las etapas vitales que nos han acompañado todos estos años, y las que nos quedan por compartir. Gracias de corazón.

Gracias a todas/os las/os compañeras/os del equipo Evaluación, Clínica y Salud, por el apoyo recibido y por los ratos de desconexión que tanto necesitaba. Gracias, también, a las que acabaron siendo amigas para siempre.

Gracias a mis compañeras de Deusto Psych por la calidad (y calidez) humana, por el cariño que me brindáis siempre. Por sostenerme cuando lo he necesitado.

Gracias a Jose Juan Uriarte y a Conchi Moreno por darme la posibilidad de desarrollar este proyecto. Por confiarme la validación de un instrumento que es, de alguna manera, especial. Gracias también a los profesionales de la Red de Salud Mental de Bizkaia que han hecho posible este proyecto, y a todos los usuarios que son parte de él.

Gracias a la Universidad de Deusto por haber colaborado en la financiación de este proyecto de investigación.

Por último, esta tesis se la dedico a Sara, mi pequeño motor, mi hija. Gracias por la paz, luz y amor que me das cada día. El tiempo y esfuerzo invertidos habrán valido la pena si con ello te inspiro al menos un poquito en perseguir todo lo que te propongas en la vida.

A todos vosotros, mi eterno agradecimiento.

Contenido

AGRADECIMIENTOS	3
ÍNDICE DE TABLAS.....	9
ÍNDICE DE FIGURAS	10
RESUMEN.....	11
ESTRUCTURACIÓN DE LA TESIS.....	13
JUSTIFICACIÓN PERSONAL DE LA TESIS	17
CAPÍTULO 1	21
EVALUACIÓN DE RESULTADOS EN SALUD MENTAL	21
1.1.- Evaluación de Resultados – Outcome Assessment (ER-OA).....	21
1.2.- La Evaluación Rutinaria de Resultados (ROM) en Salud Mental.....	26
1.2.1.- Pasos en la implementación de la ER-OA	27
1.2.2.- Medidas o Indicadores de Resultado	29
1.3.- Evaluando el cambio-protocolos de evaluación-toma de decisiones clínicas	31
1.3.1.- Evaluar el Cambio Fiable en el Individuo	32
1.3.2.- Uso de Medidas Estandarizadas para Evaluar los Resultados	34
1.3.3.- Instrumentos de Evaluación de Resultados de Servicios de Salud Mental	35
1.4.- Monitorización Rutinaria de Resultados (ROM) a nivel Internacional ...	37
1.4.1.- ROM en Salud Mental en España, País Vasco, y Bizkaia	40
CAPÍTULO 2	45
HoNOS - HEALTH OF THE NATION OUTCOME SCALE	45
2.1.- La escala HoNOS	45
2.1.1.- Génesis y Desarrollo	45
2.1.2.- Descripción y características de aplicación.....	46
2.1.3.- Estructura Dimensional del HoNOS	48
2.1.4.- Versiones del HoNOS	50

2.2.- Uso internacional y vinculación a servicios de salud	51
2.2.1.- Resultados a nivel internacional.....	52
CAPÍTULO 3	55
REVISIÓN SISTEMÁTICA DE LAS EVIDENCIAS DE VALIDEZ Y FIABILIDAD DE LA ESCALA HONOS: ADAPTACIONES Y VALIDACIONES INTERNACIONALES DE LA ESCALA.....	55
3.1.- Método	56
3.1.1.- Diseño.....	56
3.1.2.- Calidad metodológica	57
3.1.3.- Integración meta-analítica de datos clínicos	59
3.2.- Resultados.....	60
3.2.1.- Resultados de la estrategia de búsqueda	60
3.2.2.- Validez de contenido / aparente	62
3.2.3.- Validez estructural o factorial	62
3.2.4.- Validez de criterio / convergente	64
3.2.5.- Validez discriminante.....	65
3.2.6.- Validez predictiva	65
3.2.7.- Validez intercultural	66
3.2.8.- Estudios de adaptación y/o traducción a nivel internacional.....	66
3.2.9.- Sensibilidad al cambio.....	67
3.2.10.- Fiabilidad	67
3.2.11.- Interpretabilidad.....	69
3.2.12.- Puntos de corte y Error estándar de medida.....	69
3.3.- Discusión	71
CAPÍTULO 4	75
OBJETIVOS.....	75
4.1.- Objetivo(s)	75
CAPÍTULO 5	79
ESTUDIO 1: EVIDENCIAS DE ADECUACIÓN PSICOMÉTRICA DE LA VERSIÓN ESPAÑOLA DE LA ESCALA HONOS.....	79
5.1.- Objetivo.....	79
5.2.- Método	79
5.2.1.- Contexto de realización de los estudios.....	79

5.2.2.- Muestra	80
5.2.3.- Instrumentos	81
5.2.4.- Procedimiento	82
5.2.5.- Aspectos éticos	82
5.2.6.- Análisis de datos	83
5.3.- Resultados	85
CAPÍTULO 6	89
ESTUDIO 2: NUEVAS EVIDENCIAS DE ADECUACIÓN PSICOMÉTRICA DE LA VERSIÓN ESPAÑOLA DEL HONOS	89
6.1.- Objetivo	89
6.2.- Método	89
6.2.1.- Muestra	89
6.2.2.- Instrumento	91
6.2.3.- Procedimiento	91
6.2.4.- Aspectos éticos	91
6.2.5.- Análisis de datos	92
6.3.- Resultados	92
CAPÍTULO 7	101
ESTUDIO 3: BAREMACIÓN DE LA ESCALA HONOS EN POBLACIÓN BIZKAÍNA	101
7.1.- Objetivo	101
7.2.- Justificación	101
7.3.- Método	102
7.3.1.-Análisis estadísticos	102
7.4.- Resultados	107
CAPÍTULO 8	113
DISCUSIÓN GENERAL	113
8.1.- Hallazgos principales	113
8.2.- Aspectos destacados	117
8.3.- Limitaciones	120

8.4.- Líneas futuras.....	121
8.5.- Líneas de aplicación	122
CAPÍTULO 9	125
CONCLUSIONES	125
REFERENCIAS.....	1329
Anexo 1: Escala HoNOS - Health of the Nation Outcome Scale	163
Glosario de la Hoja de Puntuación HoNOS	163
1. Conducta hiperactiva, agresiva, disruptiva o agitada	163
2. Autolesiones no accidentales	164
3. Consumo Problemático de Alcohol o Drogas.....	164
4. Problemas Cognitivos.....	165
5. Problemas por enfermedad física o discapacidad.....	165
6. Problemas asociados a la presencia de ideas delirantes y alucinaciones.....	166
7. Problemas en relación con el humor depresivo	166
8. Otros problemas mentales o conductuales	167
9. Problemas con las relaciones.....	167
10. Problemas con las actividades de la vida cotidiana	167
11. Problemas con las condiciones de vida	168
12. Problemas en relación con la ocupación y las actividades	169
Anexo 2: Hoja de Puntuación HoNOS	170
Anexo 3: Clasificación Artículos Revisión Sistemática	172
Anexo 4: Clasificación Artículos Psicométricos Revisión Sistemática.....	189
Anexo 5: Hoja informativa del proyecto.....	200
Anexo 6: Consentimiento informado para el usuario participante	201
Anexo 7: Hoja informativa del Proyecto para tutor legal	202
Anexo 8: Consentimiento informado en caso de tutor legal	203
Anexo 9: Evidencias de Adecuación Psicométrica de la Versión Española de la Escala HoNOS	204

ÍNDICE DE TABLAS

Tabla 1.1.- Medidas de Resultado más utilizadas en Salud Mental (Revisión de Gilbody et al, 2003).....	36
Tabla 2.1.- Estructuras factoriales sugeridas para el HoNOS	49
Tabla 3.1.- Estructuras factoriales sugeridas para el HoNOS	63
Tabla 3.2.- Integración validez convergente instrumentos más usados	64
Tabla 3.3.- Integración meta analítica de los datos de fiabilidad recogidos.....	68
Tabla 3.4.- Integración de datos clínicos por países y para el total de estudios revisados	70
Tabla 5.1.- Estadísticos descriptivos y medidas de tendencia central de los ítems y dimensiones del HoNOS	84
Tabla 5.2.- Índices de los Análisis Factoriales Confirmatorios de los distintos modelos estructurales (índices robustos)	86
Tabla 5.3.- Validez concurrente	87
Tabla 5.4.- Diferencias por tipo de centro asistencial en las puntuaciones HoNOS	87
Tabla 6.1.- Características sociodemográficas de la muestra.....	90
Tabla 6.2.- Estadísticos descriptivos y medidas de tendencia central de los ítems y dimensiones del HoNOS (n=3715)	93
Tabla 6.3.- Datos descriptivos y t de Student para la diferencia de medias entre Incidentes y Prevalentes.....	95
Tabla 6.4.- Escalabilidad de Mokken (Hi) por escala y Samejima's graded response model parameters	96
Tabla 6.5.- Índices de Ajuste de los modelos para la muestra total y las submuestras (Incidentes y Prevalentes).....	98
Tabla 7.1.- Datos descriptivos y t de Student para la diferencia de medias entre Incidentes y Prevalentes	107
Tabla 7.2.- Baremos para la escala total de HoNOS para la muestra total, muestra incidentes y prevalentes y para hombres y mujeres según el tipo de recurso asistencial	108
Tabla 7.3.- Baremos para la subescala problemas sociales de HoNOS para la muestra total, muestra incidentes y prevalentes y para hombres y mujeres según el tipo de recurso asistencial	109
Tabla 7.4.- Puntos de Corte y dif, mínima total HoNOS y subescala social.....	110

ÍNDICE DE FIGURAS

Figura 3.1.- Diagrama de Flujo. Revisión Sistemática sobre el HoNOS	61
Figura 3.2.- Validez convergente GAF, BPRS, CANSAS y SCAN	65
Figura 3.3.- Integración meta-analítica de valores alpha del HoNOS.....	69
Figura 6.1.- Curvas características de los ítems en muestra incidente-prevalente	97

Resumen

La evaluación de resultados rutinarios (Routine Outcome Assessment – ROA) está ampliamente implementada en los servicios de salud mental a nivel internacional, donde destacan Reino Unido y Australia. Este sistema se presenta como una alternativa de elección para la valoración de la efectividad de los tratamientos aplicados. Para su implementación se hace necesario contar con instrumentos de medida adecuadamente adaptados y validados al contexto de aplicación. Uno de estos instrumentos es la escala Health of the Nation Outcome Scales – HoNOS. En las últimas décadas, se evidencia en España la necesidad de implementar este sistema en los servicios de salud. Por lo tanto, esta tesis recoge, en un primer apartado, las evidencias de validez y fiabilidad de la escala HoNOS a nivel internacional, realizando una integración meta-analítica de los datos encontrados, y ofrece a su vez los puntos de referencia clínica de la misma. En un segundo apartado y como propósito principal de esta tesis, se muestran las propiedades psicométricas y la validez estructural de la escala en una muestra reducida de 281 usuarios a través de un estudio piloto, por un lado. Por otro lado, se recogen los datos de la población de la Red de Salud Mental de Bizkaia entre los años 2013-2016 hasta noviembre de 2020 y se muestran las propiedades psicométricas de la escala HoNOS según la teoría clásica de los test (TCT) y la teoría de respuesta al ítem (TRI). Además, esta tesis ofrece los baremos de la escala HoNOS en población vizcaína, así como los puntos de corte y valores de referencia de cambio clínico a través de la Puntuación Fiable de Cambio (PFC). En conclusión, esta tesis evidencia la adecuación psicométrica de la versión española de la escala HoNOS como una herramienta clínica para la valoración de la severidad global de los usuarios.

Estructuración de la Tesis

La escala *HoNOS – Health of the Nation Outcome Scale* (Wing, Beevor, Curtis, Park, Hadden y Burns, 1998) evalúa la gravedad clínica en personas con enfermedad mental grave, y es uno de los instrumentos más utilizados para la evaluación de resultados de forma rutinaria en los servicios de salud mental.

La presente tesis tiene como objetivo principal la validación en población española de la escala HoNOS en su versión al castellano, traducida en 1999 por Uriarte, Beramendi, Medrano, Wing, Beevor y Curtis.

En la primera parte, la introducción, se encuentran cuatro capítulos, tres de ellos recogen los conceptos teóricos en los que se sustenta el trabajo de investigación y a partir de los cuales esta tesis se desarrolla. El cuarto capítulo expone los objetivos de investigación general y específicos.

En el primer capítulo se desarrolla la Evaluación de Resultados en Salud (ERS), los pasos necesarios para implementarla en servicios de salud mental y los indicadores de resultado e instrumentos de medida más empleados para poder realizarla. También se expone su implicación en la toma de decisiones clínicas y cómo poder evaluar el Cambio Fiable en el Individuo. Por último, el capítulo recoge también el estado del arte y el recorrido de la ERS, a nivel internacional primero, para después analizar su recorrido en España, País Vasco y Bizkaia, hasta llegar a la Red de Salud mental de Bizkaia (RSMB).

El capítulo dos hace un análisis de la escala HoNOS, objeto de esta tesis. Expone el contexto de su desarrollo, características de aplicación, estructura dimensional y las diferentes versiones que han surgido a raíz de la escala original, también conocida como familia del HoNOS (HoNOS Family of Measures). Además, se analiza su amplio uso y vinculación en servicios de salud mental a nivel internacional.

A continuación, en el capítulo tres, se presenta un estudio de revisión sistemática, realizada con el objetivo de recoger las evidencias de adecuación psicométrica de la HoNOS

a nivel internacional desde su desarrollo en 1998 hasta la actualidad. En él se analizan las cualidades métricas en base al listado COSMIN y se realizan las síntesis de las evidencias de fiabilidad, validez y sensibilidad al cambio. Además, a través de una integración meta-analítica de datos clínicos, se ofrecen valores de Cambio Mínimo Significativo (CMS) y Puntos de Corte (PC) calculados a partir de las fórmulas de Jacobson y Truax (1991). En conjunto, los resultados obtenidos muestran un comportamiento psicométrico adecuado de la escala para su utilización en la valoración y monitorización en procesos de enfermedad mental grave. Por otra parte, se constata que la escala HoNOS es un instrumento clínico de referencia en España pese a no contar con estudios de adecuación psicométrica, por lo que vemos necesarios su validación en muestra española.

En el capítulo cuatro se expone el objetivo general del estudio: obtener datos del comportamiento psicométrico de la escala HoNOS en población española. A su vez, se exponen tres objetivos específicos: 1) ofrecer evidencias de adecuación psicométrica en un estudio piloto, 2) ofrecer nuevas evidencias de adecuación psicométrica en población de usuarios de la RSMB, y 3) ofrecer baremos y valores de referencia de cambio clínico y puntos de corte para dar soporte a los profesionales de la salud mental en la toma de decisiones clínicas.

La segunda parte de la tesis hace referencia a los estudios realizados, recogiendo la metodología y resultados propios de cada uno. El capítulo quinto presenta el estudio piloto, el cual responde al primer objetivo específico. Se detalla la metodología y las características de la muestra que lo conforma, y se evidencia la adecuación psicométrica de la escala HoNOS.

El capítulo sexto recoge el segundo estudio de investigación, el cual responde al segundo objetivo específico. En este capítulo se describe el procedimiento y metodología del estudio de población y describe las características de la misma. Además, se reúnen las evidencias de fiabilidad y validez de la escala HoNOS obtenidas a través de la teoría Clásica de los Test (TCT) y de la Teoría de Respuesta al Ítem (TRI).

Por último, el séptimo capítulo responde al tercer objetivo específico. En este estudio se recogen los baremos que servirán para la clasificación y evaluación de usuarios del Programa de Enfermedad Mental Grave (PEMG), así como los puntos de corte y valores de referencia de cambio clínico a través de la Puntuación Fiable de Cambio en función del sexo y de los servicios de atención clínica. Los baremos permiten conocer la posición relativa del individuo

respecto a su grupo de referencia. Por su parte, el Índice de Cambio Fiable permite conocer si el cambio que realiza el usuario es estadísticamente significativo y clínicamente relevante.

Finalmente, en el tercer y último bloque se recogen dos capítulos. El capítulo de discusión (capítulo 8), donde se recoge la discusión general de este trabajo de investigación, además de exponer las posibles implicaciones prácticas de los resultados encontrados, las limitaciones del estudio y las posibles líneas futuras. Y el capítulo 9, donde se recogen las conclusiones más relevantes.

JUSTIFICACIÓN PERSONAL DE LA TESIS

La oportunidad de realizar este trabajo de investigación surge durante el desarrollo de investigaciones previas, las cuales tenían como finalidad identificar los instrumentos clínicos más usados en España y analizarlos para ofrecer valores de referencia de cambio clínico y puntos de corte. Esta línea de trabajo parte de un contexto terapéutico, con el objetivo de contar con herramientas estadísticas que sustenten la toma de decisiones clínicas.

En ese tiempo, surgió la posibilidad de colaborar en 2017 con la Red de Salud Mental de Bizkaia para realizar la validación de un instrumento que la Red aplicaba en los recursos asistenciales desde hacía muchos años, la escala HoNOS. Incluida en el protocolo de valoración inicial, esta escala fue traducida por Jose Juan Uriarte (jefe clínico de la RSMB) y su equipo, en colaboración con los autores originales en 1999. Este proyecto, el cual surge gracias a una colaboración previa, cubre la necesidad de validar psicométricamente la escala en población española. Se ofrecen también los baremos y valores de referencia de cambio clínico, herramientas prácticas para el día a día de los profesionales clínicos de la Red.

La escala HoNOS tiene un recorrido clínico de casi 20 años en la Red. El propósito final de esta tesis es ofrecer a los profesionales, no solo el uso clínico de la escala HoNOS, sino también una herramienta metodológica que apoye el trabajo en su día a día a través de la evaluación de resultados en salud.

INTRODUCCIÓN

Capítulo 1

Evaluación de Resultados en Salud Mental

La evaluación de resultados en salud mental, al igual que en otras disciplinas sanitarias, se ha desarrollado de forma creciente como opción válida en la valoración de la efectividad de las intervenciones clínicas y como alternativa a la evaluación de la eficacia que ofrecen los ensayos clínicos (Slade, 2002; Schmitter-Edgecombe et al., 2012). En el contexto natural de intervención clínica la utilización de procedimientos de investigación controlados es una opción de difícil implementación, de ahí que la metodología desarrollada por la evaluación de resultados – basada en un procedimiento de aplicación sistemático que permita la estimación del cambio (comparaciones intra y entre grupos), la utilización de instrumentos de medida adecuadamente ajustados y un análisis correcto de los datos– sea una opción cada vez más utilizada (Wise, 2004; Walton et al., 2015).

1.1.- Evaluación de Resultados – Outcome Assessment (ER-OA)

La evaluación de resultados en el ámbito de la salud en general, y de la salud mental en particular, surge como parte de un enfoque más amplio hacia la evaluación continua y la mejora de la calidad en los servicios de salud. En 1987 se publicó el artículo “evaluación de resultados 70 años después: ¿estamos preparados?” [Outcome assessment 70 years later: are we ready?] (Shroeder, 1987) que aludía a los intentos de conseguir que cirujanos y hospitales publicaran los resultados de las operaciones a principio del siglo XX. Por entonces, Ernst Codman, un cirujano estadounidense, sostuvo que los hospitales debían informar sobre los resultados de sus tratamientos de forma armonizada, para permitir comparaciones entre ellos (Codman, 1914). Además, sugirió que todos los aspectos de la asistencia hospitalaria deberían ser evaluados para garantizar que producían efectos favorables (Jefford, Stockler y Tattersall, 2003).

A medida que la atención médica evoluciona, se reconoce la necesidad de medir y analizar los resultados de las intervenciones para comprender mejor la eficacia de los tratamientos y mejorar la atención proporcionada. Donabedian en 1966 consagró el término “resultado” (*outcome*, en inglés) a través de sus decisivas contribuciones al análisis y medida de la calidad de los servicios sanitarios. En particular, formuló la trilogía “estructura”, “proceso” y “resultado” como un nuevo paradigma para la evaluación de la calidad (Donabedian, 1966). Planteó que ciertos resultados, como la muerte, son fácilmente reconocibles y medibles, mientras que otros eran difíciles de definir y medir. Entre estos últimos incluía las actitudes y satisfacción de los pacientes, la reinserción social, la discapacidad física y la rehabilitación. Sin embargo, destacó estos resultados de salud como parte fundamental para poder validar la efectividad y la calidad de la atención médica proporcionada (Jefford, Stockler y Tattersall, 2003). Este enfoque se alinea con la creciente importancia de la medicina basada en la evidencia y la atención centrada en el paciente.

Las primeras publicaciones las encontramos en los años 80. McPheeters (1980) discutía el uso de algunas escalas de uso común para evaluar el impacto comunitario de los servicios de salud mental. Poco después, Ciarlo (1982) escribía sobre la llegada de la evaluación de los resultados a nivel de usuario buscando evidencias de resultados positivos o impacto en los clientes para justificar la implementación de programas y su mantenimiento. La evaluación de resultados ofrece una forma de contabilización a través de la evaluación del rendimiento y garantía de calidad (Trauer, 2010).

La evaluación de resultados de los servicios de salud permite alcanzar el objetivo de mejorar los tratamientos ofrecidos (National Institute for Mental Health [NIMH], 2008). Promueve la búsqueda de una mejora continua al proporcionar información clave para la mejora de los servicios de salud mental, identificando las áreas con buen funcionamiento respecto a otras que necesitan un mayor desarrollo (Jacobs, 2007). Además, la evaluación periódica del paciente permite tomar decisiones sobre la idoneidad del tratamiento en base a la efectividad de las intervenciones y su continuación, cambio o finalización (Kazdin, 2008). Por otra parte, la información recogida permite a los profesionales y a las organizaciones involucradas mostrar la eficacia de los servicios ofrecidos, al mismo tiempo que impulsa la investigación clínica al tratar datos relevantes recogidos de los usuarios.

Pero ¿qué es un resultado? Más allá de considerarlo como un efecto o producto, en salud mental la idea de un resultado es entendida como la consecuencia de una acción, a los cambios observables o medibles en la salud y el bienestar de un individuo como consecuencia de la

atención recibida. Estos cambios incidirían, entre otros aspectos, en la funcionalidad, en los síntomas y otros elementos relevantes para la salud mental. Esto se refleja en la definición de resultado del grupo de Andrews, quienes lo consideran como el efecto en el estado de salud de un paciente y que se puede atribuir a una intervención (Andrews et al., 1994).

Puede parecer que el requisito mínimo para un resultado es que algo haya cambiado. Sin embargo, se ha señalado que "mantener el estado de salud de un paciente puede considerarse en algunas circunstancias un resultado positivo" (Jacobs, 2009, p. 3). Esto es especialmente pertinente en el caso de los pacientes con condiciones crónicas, o incluso deterioradas, de los que no se espera de forma realista que obtengan mejoras significativas. Incluso cuando se obtienen cambios en las evaluaciones, puede haber problemas técnicos para juzgar si un cambio en la puntuación obtenida por un indicador o instrumento se corresponde con un cambio en el paciente. Es necesario juzgar si un cambio en la puntuación del indicador es puramente aleatorio, una imprecisión en la medición, una fluctuación clínicamente trivial o un cambio verdadero y significativo en el estado de salud. Una perspectiva útil sobre la centralidad del cambio fue propuesta por Eagar (2002), quien distinguió entre "antes y después" y "con y sin". El primero ("antes y después") modela el resultado como el cambio en las puntuaciones debidas a una intervención entre el momento anterior y posterior a su implementación. Pero, como afirma Eagar, el enfoque "con y sin" es "de especial relevancia para las enfermedades crónicas en las que el objetivo de la intervención puede ser mantener el estado de salud actual, el cual se calificaría como un buen resultado si la alternativa fuera un posible deterioro" (Trauer, 2010; pp. 143).

La evaluación de resultados en salud es aquella actividad que utiliza métodos de investigación experimentales o naturalísticos (muchas veces observacionales) para medir los resultados de las intervenciones sanitarias en condiciones de práctica clínica habitual (Moses, 1995; D'Agostino & Kwan, 1995; Sacristán, Galende & Soto, 1998). Se examinan las consecuencias derivadas de las intervenciones, identifica los determinantes de las diferencias entre eficacia y efectividad, y una vez evaluada la eficacia, la evaluación de resultados en salud permite analizar los resultados obtenidos en condiciones de práctica clínica habitual (Badia y Bigorra, 2000), esto es, su efectividad y eficiencia.

Hay cinco características generales que definen la evaluación de resultados en salud (Badia y Bigorra, 2000):

- 1- Está basada en la práctica clínica habitual.

- 2- Se basan en el análisis de la efectividad de las intervenciones de salud. Se interesan en comprender y analizar efectos sinérgicos con la intervención de otros factores de influencia.
- 3- Énfasis en los beneficios del paciente. Prestan mayor atención a variables importantes para el propio paciente. Esta perspectiva es esencial, ya que los resultados deben reflejar las metas y valores individuales. La participación del paciente en el proceso de evaluación garantiza que los objetivos terapéuticos estén más alineados con sus necesidades y expectativas. Es imprescindible la utilización de una mayor diversidad de variables de resultado, especialmente las que resultan más relevantes para el propio paciente.
- 4- Se utilizan métodos de investigación establecidos. Énfasis en analizar las condiciones de práctica clínica habitual.
- 5- Se prioriza un enfoque individual (paciente) y/o agregado (poblacional). Los resultados obtenidos pueden reflejar las respuestas de grupos agregados. Especial empeño en los resultados que se circunscriben al paciente individual. Importa conocer cómo el paciente particular se presenta al inicio de la intervención, su evolución y los cambios experimentados como resultado del tratamiento. Interesa conocer cuál es la situación de cambio respecto a si mismo como respecto a otros grupos de referencia cuya comparación puede basarse en estándares de salud de grupos clínicos como de población general.

Leginski et al., (1989) desarrollaron la paráfrasis ‘quién recibe qué de quién, a qué coste y con qué efecto’, para identificar los agentes y componentes comprometidos en la evaluación de resultados. El quién se refiere a los usuarios. El qué se refiere a los servicios prestados. El de quién se refiere al personal. A qué coste es, obviamente, el presupuesto disponible. Con qué efecto, se trata de los resultados medidos.

De forma análoga, Iraurgi (2010), inspirado en un verso de Kipling (1901)¹, propone la paráfrasis: “Qué hacemos a Quién para obtener Qué y en Dónde”, análoga al acrónimo PICO (Patients, Intervention, Comparative & Outcomes), tratando de identificar todas las variables implicadas en una pregunta de investigación: la variable independiente (qué hacemos), la

¹ “Mantengo seis honrados sirvientes (ellos me enseñaron todo lo que sé); sus nombres son: Qué y Por qué, y Cuándo y Cómo, y Dónde y Quién”. Kipling, R (1901). *Just so Stories*. MacMillan.

unidad de análisis (a quién), la variable dependiente (para obtener qué), y las variables contextuales o terceras variables de influencia (en dónde).’

Todo ello muestra la relación de los ‘resultados’ con otros aspectos fundamentales de los servicios humanos complejos. Alerta sobre la necesidad de tener en cuenta a los pacientes, los servicios, las finanzas y el personal a la hora de evaluar el rendimiento de un servicio.

En otro ámbito de aproximación en los resultados de salud, el empoderamiento confiere a las personas con experiencia propia una mayor comprensión y un mayor control sobre su vida (OMS, 2010). Requiere que los gobiernos, los empleadores, las instituciones educativas, las organizaciones no gubernamentales y el público en general eliminen las barreras que pueden obstaculizar la participación plena y real en la sociedad de las personas con experiencia propia de problemas de salud mental. En las últimas décadas ha habido un cambio progresivo hacia una mayor participación de los usuarios de los servicios y sus cuidadores en diferentes niveles del sistema de atención de la salud mental:

- Nivel personal: participación en la planificación, la evaluación y la gestión de la propia atención de salud, por ejemplo, a través la toma de decisiones compartida, la planificación anticipada, el apoyo a la autogestión y los enfoques de atención orientados a la recuperación y centrados en las personas.
- Nivel comunitario: participación en la planificación de la prestación, el seguimiento y la evaluación de los servicios locales, la promoción, las campañas de concientización (especialmente las destinadas a combatir la estigmatización), la capacitación del personal que trabaja en la salud mental y otros.
- Nivel estratégico: participación en la formulación de políticas, planes y leyes de salud mental, la supervisión de los servicios y la investigación (Organización Panamericana de la Salud, 2023, pp. 120)

Los objetivos terapéuticos que alcanzar a través de indicadores serían, por un lado, lograr la estabilidad de los síntomas (Trauer, 2010). Los indicadores empleados miden la estabilidad de los síntomas a lo largo del tiempo, a la vez que proporcionan información sobre la eficacia del tratamiento aplicado. Por otro lado, estos indicadores se utilizarían a su vez para lograr objetivos funcionales, recogiendo la capacidad del usuario para realizar actividades de la vida diaria, mantener relaciones sociales o participar en comunidad. Esta evaluación de resultados contribuye de forma significativa a la calidad de la atención y al bienestar de los individuos atendidos.

En todos los casos, identificar e involucrar a las partes interesadas en la formulación de leyes y políticas de salud mental es importante para garantizar que sean adecuadas para su propósito. Los planes detallados en los que se describa cómo se aplicarán las leyes y políticas son vitales. Esto implica:

- establecer plazos para elaborar y entregar elementos específicos de la ley y las políticas;
- asignar presupuesto;
- estimar las necesidades de recursos humanos (tanto especializados como no especializados) y elaborar un plan sobre cómo se capacitarán y ubicarán;
- determinar todas las actividades que se deben realizar y establecer funciones y responsabilidades claras para cada una, y
- establecer cómo se realizará el seguimiento y la evaluación para analizar el progreso y el desempeño (incluido el cumplimiento de los derechos), y para favorecer la mejora continua (Organización Panamericana de la Salud, 2023).

1.2.- La Evaluación Rutinaria de Resultados (ROM) en Salud Mental

La Evaluación de Resultados de Forma Rutinaria (Routine Outcome Monitoring o ROM) en salud mental implica la recopilación sistemática de datos sobre el progreso del paciente a lo largo del tiempo utilizando indicadores, instrumentos y procesos estandarizados (De Jong, et al., 2022). La información obtenida permite al profesional y al paciente la posibilidad de realizar un seguimiento durante toda la duración del tratamiento (Lambert, et al., 2001).

Numerosos estudios, así como la propia práctica clínica, sugieren claros beneficios en la evaluación de resultados en las terapias psicológicas (Lambert et al., 2018; La Tona et al., 2023). Esta evaluación puede ayudar a los profesionales y a los pacientes a identificar posibles dificultades en áreas específicas, las cuales podrían pasar desapercibidas sin este procedimiento (Lambert et al., 2018).

Slade (2002) publicó en una revisión sistemática cuales son los siete dominios clave que se han utilizado para la evaluación de resultados en salud mental. Estos incluyen el bienestar (o calidad de vida), el funcionamiento cognitivo o emocional, el comportamiento

(behavioural activities), la salud física (y la discapacidad), el funcionamiento interpersonal y social, la sociedad (la carga de los cuidados y la seguridad pública) y la satisfacción con los servicios.

1.2.1.- Pasos en la implementación de la ER-OA

En ese mismo estudio, Slade (2002) describe cómo introducir la evaluación de resultados a través de cuatro pasos.

Paso 1: Valorar los recursos clave necesarios

Es importante considerar las propiedades psicométricas de cualquier medida de resultado para que la implementación de la ER-AO sea útil. Deben establecerse cuáles son las propiedades adecuadas y poder administrarse de forma estándar. Todos los datos recogidos deben servir para el tratamiento de los pacientes o para la evaluación de los servicios. Para esto, se necesita un sistema para el análisis de los datos y obtener, así, feedback del proceso.

Por otra parte, considerar este primer punto va a permitir identificar cuáles son los recursos necesarios, como capacidad de liderazgo, experiencia, contar con personal de apoyo, tecnología de la información (Ej. acceso a ordenadores, software fácil de usar) y tiempo clínico. En caso de no disponer de estos recursos, no sería posible implementar la evaluación rutinaria de resultados, ya que podría dar lugar a una recogida de datos de baja calidad, sin utilidad ni capacidad de feedback.

Uno de los recursos más valiosos a considerar es el tiempo disponible del profesional. En la práctica clínica habitual, la medición y el análisis de resultados puede suponer un aumento del 10% del tiempo que dedica el clínico a cada paciente (Marks, 1998). A principios de los 2000, los clínicos no estaban convencidos de que este tiempo extra (y, en consecuencia, la reducción del número de pacientes que pueden atender) merezca la pena. Dado que la atención clínica se considera un tiempo valioso y rellenar formularios no se ve como "trabajo", existía una fuerte desmotivación para cumplimentar las medidas de resultados (Slade, 1999).

En los últimos años, la cultura en la atención clínica en algunos servicios de salud mental ha cambiado, ya que se han elaborado protocolos de evaluación que funcionan a nivel nacional (UK, Penington, et al., 2022; Australia, Burgess, et al., 2015).

La evaluación de resultados se implementa con éxito cuando los profesionales clínicos valoran la relevancia y utilidad clínica de la propia evaluación. Los clínicos deben sentir como propios los datos obtenidos para comprometerse con su uso y apreciar los beneficios que proporcionan.

Tradicionalmente, los clínicos consideran que su responsabilidad es para con el paciente, mientras que el uso de medidas de resultados suele requerir una perspectiva más grupal o poblacional (Hall, 1996; Callaly y Hallebone, 2001). Por lo tanto, se hace necesario un cambio en la cultura, para que la atención pase de centrarse en las medidas de proceso (rendimiento) a centrarse en las medidas de resultados. Para ello, podría ser necesario limitar el número de casos y así garantizar un nivel definido de calidad asistencial. Se podrían considerar otras opciones, como la introducción de incentivos de pago para los clínicos, el seguimiento de los datos de resultados durante la formación y otras formas de cambiar la cultura (Slade, et al., 1999).

Será necesario, también, encontrar recursos creativos para minimizar el tiempo que el personal clínico dedica a la recogida y el análisis de datos. Por ejemplo, la introducción automatizada de datos, dar acceso del recurso a personal no clínico, el uso de equipamientos multi-media para proporcionar información instantánea (utilizando datos introducidos previamente para trazar el progreso a lo largo del tiempo) sería una forma de minimizar la carga y maximizar los beneficios potenciales para la atención clínica.

Por tanto, en este primer paso es necesario hacer un cuidadoso inventario de todos los recursos necesarios para implantar el sistema, asegurando la disponibilidad de todos ellos para facilitar un proceso sencillo de recogida de datos y feedback de los mismos, de modo que puedan apreciarse fácilmente los beneficios y se produzca un cambio en la cultura.

Paso 2: Elegir los dominios a evaluar

Se trata de identificar los ámbitos de resultados que conviene medir en el servicio. Éstos variarán entre dos extremos que comprenden lo que el paciente considera importante (su experiencia individual) hasta lo que el clínico define como importante (comprender las experiencias mentales anormales). Un servicio que atiende la perspectiva del paciente estará más interesado en los dominios de resultados relacionados con el cambio experiencial individual. Un servicio que recoge el interés del profesional estará más interesado en los dominios de resultados relacionados con los síntomas y el funcionamiento, y en garantizar que las intervenciones aplicadas estén de acuerdo con la evidencia.

Un planteamiento para aunar los intereses de ambas partes propone evaluar los siete ámbitos mencionados anteriormente, lo cual puede resultar poco práctico. Otro enfoque plantea evaluar tan solo un ámbito, lo cual podría no ser suficientemente significativo. La mayoría de los enfoques, para implementar la evaluación rutinaria de resultados, recomiendan evaluar entre dos y cuatro dominios.

Paso 3: Detalles técnicos y cómo se evaluarán los dominios escogidos

Para elegir la forma de evaluación de los dominios relevantes, es necesario atender a las siguientes cuestiones: ¿Qué constituye un buen resultado para un paciente que no contempla recuperarse?, ¿El objetivo es demostrar que el tratamiento causó la mejoría o sólo que ésta se produjo sin referencia al tratamiento?, ¿Se centra el resultado sólo en el paciente, o también en sus familiares o cuidadores?, ¿Se centra en el cambio individual o en el conjunto de los datos para examinar los cambios a nivel grupal?, ¿Se prefieren medidas directas del paciente o del cuidador, o indirectas por el personal o uso del servicio?, ¿Se prefieren medidas genéricas (aplicables a grupos amplios) o medidas específicas (para afecciones concretas)?, ¿Se prefieren medidas individualizadas o estandarizadas?, ¿Debe realizarse la evaluación en momentos "importantes" del proceso asistencial o a intervalos fijos?, ¿Cómo deben interpretarse los cambios "contradictorios", como el aumento de los síntomas acompañado de un aumento de la calidad de vida? Las respuestas a estas preguntas determinarán los principios clave del marco de trabajo en la evaluación de resultados.

Paso 4: Identificar las medidas de resultados

En este último paso se escogen las medidas de resultado que más se ajustan a los requisitos identificados en los tres pasos anteriores. Las medidas de uso rutinario no solo deben ser factibles, sino también breves, sencillas, aceptables, disponibles, pertinentes y valiosas (Marks, 1998).

Estos pasos pueden ayudar a los servicios a considerar cuidadosamente cómo implementar un enfoque bien pensado y con recursos suficientes para recopilar y utilizar datos de resultados de forma rutinaria, en lugar de algo que sea una carga administrativa añadida al trabajo "real" de los clínicos.

1.2.2.- Medidas o Indicadores de Resultado

Drapeau (2012) identificó 10 medidas, indicadores o sistemas para registrar los cambios en salud mental en la práctica clínica diaria. 1) The Celest Health System (CHS-MH); 2) The

CORE-OM (Chris Evans et al., 2000); 3) The Behavior and Symptom Identification Scale 24 (BASIS-24); 4) The COMPASS Tracking Assessment System; 5) The Outcome Questionnaire-45 (OQ-45); 6) The Outcome and Session Rating Scales (ORS and SRS); 7) Polaris-MH System; 8) PSYCHLOPS (Psychological Outcome Profiles); 9) The Schwartz Outcome Scale – 10 (SOS-10); 10) The Treatment Outcome Package (TOP)

Boswell et al. (2015), teniendo este estudio como referencia, destacó otros cuatro. PCOMS ICCE (The Partners for Change Outcome Management System: International Center for Clinical Excellence). Fue desarrollado en el año 2005 por Miller y colaboradores; Incluye medidas que cubren todos los rangos de edad, formando así un sistema de gestión de resultados que no sólo proporciona información válida y fiable, sino que también es fácil de usar para terapeutas y pacientes.

Otro sistema conocido es TOP (Treatment Outcome Package) (Youn, et al., 2012), una medida de resultados desarrollada por clínicos para su uso en entornos clínicos rutinarios; permite informar de diversos aspectos del trabajo clínico, tales como el desarrollo de formulaciones de casos, la facilitación de la comunicación cliente-terapeuta, el seguimiento de cambio durante el tratamiento (positivo y negativo) y la documentación del área específica de experiencia del terapeuta.

Por otra parte, el sistema CORE (Clinical Outcomes in Routine Evaluation) (Evans et al., 2000; Barkham, et al., 2001; 2006) ha demostrado ser una medida (CORE-OM) válida y fiable para múltiples entornos, aceptable para los usuarios y los clínicos, así como para responsables políticos de los sistemas sanitarios.

Por último, el OQ Systems (Outcome Questionnaire System) (Lambert et al., 1996). Este cuestionario fue diseñado con el objetivo de evaluar tres dominios centrales para la salud mental: malestar subjetivo, relaciones interpersonales y desempeño del rol social. El OQ muestra tener una fiabilidad alta y evidencias que sugieren una buena validez concurrente y de constructo de la puntuación total. En el estudio de validación, el cuestionario distingue entre muestras de pacientes y no pacientes, es sensible al cambio y se correlaciona con otras medidas de malestar del paciente.

A nivel mundial, los instrumentos más utilizados para el seguimiento de los resultados en el ámbito clínico son el sistema desarrollado por Miller et al. (2005) Partners for Change Outcome Management System (PCOMS; Duncan, 2012) y el Outcome Questionnaire System de Lambert et al., (1996) (La Tona et al., 2023).

1.3.- Evaluando el cambio-protocolos de evaluación-toma de decisiones clínicas

La característica central en la evaluación de resultados (ER-OA) es la capacidad de obtener mediciones en diferentes puntos temporales para detectar y medir el cambio entre esos puntos. Por sí solas, estas evaluaciones individuales solo representan medidas del estado actual (en el momento de la evaluación); tan solo se considerarán como medidas reales de resultados cuando haya dos o más evaluaciones, permitiendo valorar si ha habido un cambio en el estado o no (Trauer, 2010). Esta característica hace que la ER-OA permita dar información sobre el cambio tanto a nivel individual, relevante para médicos, profesionales de salud en atención directa, cuidadores e incluso los propios consumidores-pacientes, como a nivel grupal, información de especial interés para jefes de equipo, gestores, financiadores y políticos. De cualquier modo, ambos niveles están íntimamente conectados, ya que los datos a nivel grupal se construyen a partir de las evaluaciones que se realizan a nivel individual, por lo que muchas personas se benefician de la información obtenida en ambos niveles. Los resultados a nivel de grupo pueden informar sobre la interpretación a nivel individual, y viceversa (Cella et al., 2002). No obstante, para llegar a estos resultados se requieren métodos diferentes, ya que la confianza que proporcionan las muestras grandes de los estudios de grupo no existe a nivel individual.

Los métodos tradicionales para valorar los cambios como resultado de los tratamientos han sido las pruebas de significación estadísticas y las medidas del tamaño del efecto (Cohen, 1988). Este último método se utiliza para estimar la magnitud de la diferencia entre medidas de resultados obtenidas en distintos momentos temporales. Con el tamaño del efecto, podemos valorar si el cambio observado es suficientemente grande como para poder aceptar que tiene importancia empírica y supone un logro considerable (Iraurgi, 2009b). De igual modo, a nivel terapéutico podemos plantear que tal logro -cambio de magnitud aceptable y con baja probabilidad de haberse producido por azar- es un cambio importante para la toma de decisiones; pero ¿es clínicamente significativo? (Iraurgi, 2010).

Los contrastes estadísticos de significación han sido criticados por los clínicos, quienes argumentan que identificar diferencias estadísticamente significativas no supone necesariamente un cambio clínico relevante (Cohen, 1988; Jacobson y Truax, 1991). El valor de *p* (significación estadística) no es una medida de asociación, sino la probabilidad de que el resultado observado se deba al azar y, por ello, se ve muy influenciada por el tamaño

muestral ya que con ello se aumenta la precisión al reducirse el error de medida (Iraurgi, 2009a).

Son tres las razones por las que es impropio utilizar la significación estadística para evaluar el cambio:

- 1) La significación estadística sirve principalmente para evaluar las diferencias intergrupales (entre grupos) y no intraindividual (dentro de un individuo);
- 2) La significación estadística hace referencia a la probabilidad de que el cambio no se deba al azar, y no tanto a la magnitud de éste;
- 3) En muestras muy grandes, los resultados pueden ser estadísticamente significativos incluso si las diferencias son muy pequeñas y sin sentido clínico.

Esto ha llevado al interés en evaluar la significación clínica de los cambios observados.

1.3.1.- Evaluar el Cambio Fiable en el Individuo

Entre las diversas definiciones formales que existen sobre qué es la fiabilidad, tomaremos la que la define como la precisión con la que se mide algo. La fiabilidad puede estimarse a través de varios métodos: realizando la prueba de nuevo después de un breve intervalo (test-retest), obteniendo dos o más mediciones de diferentes evaluadores (inter-jueces) y calculando las correlaciones entre los distintos ítems dentro de la misma prueba, viendo hasta qué punto estos componentes producen los mismos valores (consistencia interna).

Independientemente del método, en todos se realizan múltiples mediciones de la misma unidad, la cual se asume invariable. Por tanto, la fiabilidad que obtenemos del instrumento surge a partir de la concordancia entre esas mediciones. La fiabilidad puede oscilar entre cero (inaceptable, muy poco fiable) y uno (excelente, totalmente fiable). Dentro de este rango, Nunnally y Bernstein (1994) sugieren que un nivel satisfactorio de fiabilidad se encontraría a partir del valor 0.70, un valor "modesto" que puede ser aceptable para la investigación preliminar y básica, mientras que cuando las mediciones son utilizadas para tomar decisiones sobre los individuos se requiere un valor de .80 o más.

Al evaluar el cambio, cuando tanto la primera puntuación como la siguiente se miden con una fiabilidad imperfecta, habrá algunas diferencias que son simplemente el resultado de esa falta de precisión. El primer trabajo importante que identificó el concepto de cambio fiable fue el de Jacobson et al. (1984), perfeccionado posteriormente por Christensen y Mendoza

(1986). La esencia de este enfoque es identificar cuán grande debe ser una diferencia entre dos puntuaciones de la misma persona para que sea muy poco probable que sea el resultado de la falta de fiabilidad, es decir, cuán grande debe ser una diferencia para estar más allá del margen de error de medición. Esta diferencia se conoce como Índice de Cambio Fiable (ICF-RCI), uno de los métodos más usados para evaluar el cambio a nivel individual (Jacobson y Truax, 1991), el cual informa si la magnitud del cambio observado excede el rango que podría ser atribuido al monto probable por error o azar.

Este método posee la ventaja de permitir el análisis del cambio a nivel individual del paciente y saber cuál es el porcentaje de las personas que cambian dentro de un grupo. Jacobson y Truax (1991) proponen dos pasos para valorar el Cambio Clínico Significativo (CCS): en primer lugar, estimar el cambio mediante el RCI y, en segundo lugar, estimar el cambio clínico o en su condición de salud que se refiere a si el sujeto ha pasado a estar más cerca de la media del grupo de población funcional que de la del grupo no funcional. Dicho de otro modo, si se desplaza del puntaje de una población clínica (disfuncional) hacia una “no-clínica” o normativa (funcional). Ambos criterios deben cumplirse para establecer que se ha producido un CCS de acuerdo con estos autores. El problema que reporta el CCS es que para poder calcularlo se debe poseer los datos de la población normativa, los cuales no siempre están disponibles.

Para que el RCI sea considerado estadísticamente significativo se toma como referencia la convención de un nivel de significación menor del 5% ($\alpha < .05$), tomando como referencia las puntuaciones z estandarizadas (Wilkinson, 1999). Debido a que solo contamos con la distribución de un grupo clínico y que el sentido de la funcionalidad se expresa en bajas puntuaciones, el 5% de los casos superiores se situarían por encima de un valor $z = 1.65$. Por lo tanto, teniendo en cuenta este criterio se calcula la Puntuación de la Diferencia Mínima de Cambio (PDMC), se obtendrá a partir del producto de $z_{\alpha} * SE_{dif}$, donde z_{α} representa el valor de probabilidad aludido más arriba, y SE_{dif} representa el Error Estándar de la Diferencia de Puntuaciones entre las medidas antes y después de la intervención².

² Un desarrollo más específico de la formulación y cálculo del RCI se expondrá en el apartado de Método del estudio 3 (Capítulo 7, pg. 101).

1.3.2.- Uso de Medidas Estandarizadas para Evaluar los Resultados

El Plan Estratégico de Investigación del NIMH (El Instituto Nacional de la Salud Mental, en sus siglas en inglés) de Estados Unidos (estrategia 4.4), exige el desarrollo de medidas válidas y fiables de la calidad del tratamiento y de los resultados que puedan aplicarse de forma factible a nivel de persona, clínica, sistema y población (Sherrill, 2015). Recientemente, el Instituto de Medicina estableció un comité ad hoc para desarrollar un marco para el desarrollo de estándares basados en la evidencia para las intervenciones psicosociales para los trastornos mentales. En su cargo y en sus deliberaciones, el comité hizo hincapié en la importancia de las medidas de calidad y de rendimiento de los resultados. Asimismo, la aplicación de la Ley de Paridad de Salud Mental y Equidad de Adicciones de 2008 y la Ley de Protección al Paciente y Cuidado de Salud Asequible de 2010 obligan a prestar mayor atención a la evaluación de la calidad y la rendición de cuentas.

La aplicación de medidas pragmáticas sobre la calidad de la psicoterapia podría facilitar la formación y la supervisión de los terapeutas (Schoenwald et al., 2011; Sherrill, 2015). Dichas aplicaciones también podrían utilizarse como herramientas para el seguimiento de la calidad y la aplicación de iniciativas de mejora de la calidad. Estas medidas, consideradas eficientes, también podrían impulsar una importante investigación en los entornos de la práctica comunitaria (por ejemplo, estudios de eficacia, implementación y mejora de la calidad). En general, existe una necesidad apremiante de contar con estrategias eficaces y eficientes para evaluar la calidad de la psicoterapia a fin de garantizar, en última instancia, la disponibilidad de intervenciones psicosociales de alta calidad.

Las medidas de fidelidad utilizadas en los estudios de eficacia tienen una utilidad limitada para su uso en la práctica comunitaria, ya que evalúan estrictamente la fidelidad de las terapias manuales específicas e implican una codificación intensiva del contenido de las sesiones. Por lo tanto, se necesitan estrategias basadas en la investigación que equilibren el rigor (fiabilidad y validez) con la viabilidad (teniendo en cuenta los recursos y las implicaciones para las rutinas clínicas y administrativas) para garantizar que los pacientes reciban intervenciones de alta calidad basadas en la evidencia (Schoenwald, et al., 2011).

Esta iniciativa pretende apoyar el desarrollo inicial de medidas pragmáticas que incorporen explícitamente las perspectivas del usuario final para garantizar que las herramientas sean viables en la práctica comunitaria; que estén fuertemente asociadas con los resultados de los pacientes; y que sean útiles para informar sobre la práctica/política (por ejemplo, las calificaciones serían útiles para supervisar la calidad e informar sobre los

esfuerzos de mejora de la calidad). La iniciativa también pretende apoyar la investigación para examinar prospectivamente la viabilidad, fiabilidad y validez de las medidas y estrategias de evaluación de la calidad resultantes, para examinar su utilidad para evaluar la calidad de la psicoterapia en los entornos de la práctica.

1.3.3.- Instrumentos de Evaluación de Resultados de Servicios de Salud Mental

Es importante seguir una serie de criterios cuando se selecciona el instrumento para la evaluación de resultados. Andrews et al., (1994) expusieron una serie de criterios que debía cumplir: deben ser aceptables y breves, y el propósito-finalidad, la redacción y la interpretación deben ser claros. Asimismo, la medida debe ser aplicable, y debe poder evaluar dimensiones que son importantes para los usuarios y los proveedores de servicios, además de proveer información que facilite el manejo de los servicios. Deben ser de bajo coste, con una mínima/sencilla formación para su aplicación, puntuación e interpretación, pero que a su vez sean válidos, fiables y sensibles al cambio, lo suficiente como para indicar que los resultados observados son debidos a la intervención (Stedman y Yellowlees, 2000).

En una revisión más reciente, Pirkis et al., (2005) sugieren once criterios según los cuales se podrían juzgar las medidas de resultado: 1) Ser relevante y apropiado, 2) Ser fácil de usar, 3) Tener significados claros y objetivos, 4) Reflejar las perspectivas de todos los participantes, 5) Ser capaz de identificar los efectos positivos del tratamiento, 6) Ser psicométricamente sólido/robusto, 7) Ser económico, 8) Ser comprensible para los no profesionales, 9) Susceptible de obtener resultados rápidos y fáciles, 10) Ser clínicamente útil, y 11) Tener una base teórica amplia.

Los instrumentos de evaluación de resultados utilizados en servicios de salud mental cumplen con los requisitos mencionados anteriormente, lo que los hace ser efectivos y fiables. Aunque no existe un "gold standard" único que sea aplicable en todos los contextos, hay varios instrumentos reconocidos y utilizados ampliamente que se expondrán a continuación. Este listado se presenta como una aproximación al conjunto de medidas más utilizadas en la evaluación de la Salud Mental, pero su elección dependerá de los objetivos específicos de la evaluación y del contexto clínico. Es importante tener en cuenta que la evaluación de resultados en salud mental puede requerir una combinación de medidas para abordar aspectos diversos del bienestar psicológico y emocional. Además, la elección del instrumento debe considerar la retroalimentación del usuario y su participación en el proceso de evaluación.

Gilbody et al., (2003) realizó una revisión sistemática en la que recogió las medidas de resultados más utilizadas en los ensayos clínicos y la investigación de resultados en psiquiatría entre los años 1955 y 2000. En este estudio descubrieron que el método dominante de medición de resultados en ensayos aleatorios en psiquiatría eran las escalas de psicopatología basadas en síntomas. Los dominios evaluados por los instrumentos eran depresión, esquizofrenia, demencia, funcionamiento social, calidad de vida y satisfacción; y también había medidas de resultados global. Algunas de las medidas de resultados más comunes se recogen en la Tabla 1.1.

Tabla 1.1.- Medidas de Resultado más utilizadas en Salud Mental (Revisión de Gilbody et al, 2003)

Instrumento de medida	Acrónimo	Área de Exploración
1. Escala de Hamilton para la depresión	HDRS	Depresión
2. Inventario de depresión de Beck	BDI	Depresión
3. Escala hospitalaria de ansiedad y depresión	HADS	Depresión
4. Escala breve de evaluación psiquiátrica	BPRS	Psicosis
5. Escala del síndrome positivo y negativo	PANSS	Psicosis
6. Mini-Mental State Examination	MMSE	Demencia
7. Escala de valoración de la enfermedad	HSRS	Medida de Resultado Global
8. Escala de evaluación global	GAS	Medida de Resultado Global
9. Escala de evaluación global del funcionamiento	GAF	Medida de Resultado Global
10. Escala de ajuste social	SAS	Funcionamiento Social
11. Índice de Katz/ Escala funcional de Katz	KFS	Funcionamiento Social
12. Escala de funcionamiento social	SFS	Funcionamiento Social
13. Índice de actividades de la vida diaria	ADL	Funcionamiento Social
14. Escala REHAB	REHAB	Funcionamiento Social
15. Análisis funcional de los entornos asistenciales	FACE	Funcionamiento Social
16. Índice de calidad de vida de Lehman	QOLI	Calidad de Vida
17. Escala de calidad de vida de Heinrichs	QLS	Calidad de Vida
18. Manchester Short Assessment of Quality of Life	MANSA	Calidad de Vida
19. Expectativas y Servicios de Cuidadores y Usuarios	CUES	Calidad de Vida

En una revisión más reciente, Burgess y colaboradores (2017) identificaron 20 instrumentos relevantes. Como referencia tuvieron dos revisiones sistemáticas previas, desarrolladas para Australia (AMHOCN, 2013) y Nueva Zelanda (Lutchman, et al., 2007; Waikato Evaluation Team, 2005). Tan solo cinco instrumentos cumplían los criterios que se consideran adecuados: (1) breve (compuesto por menos de 50 ítems) y fácil de puntuar, (2) no es redundante por instrumentos más recientes, (3) el instrumento ha sido analizado científicamente, (4) considera el funcionamiento desde una perspectiva contemporánea / actualizada, y (5) demuestra propiedades psicométricas sólidas. Los cinco instrumentos fueron: Life Skills Profile (LSP-16), the Health of the Nation Outcome Scales (HoNOS), the Illness Management and Recovery Scale–Clinician Version (IMRS-C), the Multnomah Community Ability Scale (MCAS) and the Personal and Social Performance Scale (PSP).

1.4.- Monitorización Rutinaria de Resultados (ROM) a nivel Internacional

La evaluación rutinaria de los resultados (ROM – Routine Outcome Monitoring) se ha llevado a cabo de diversas maneras a nivel internacional. Los enfoques utilizados por los países que han implementado la ROM tienen como objetivo informar sobre programas y sistemas, en lugar de ayudar a informar sobre los resultados del tratamiento a nivel de paciente individual.

En diciembre de 2002, se estableció en Reino Unido un programa nacional para dirigir la introducción de la medición sistemática de los resultados en salud mental. Tras una prueba piloto, el Profesor Louis Appleby, director nacional de Salud Mental, conformó el Grupo de Referencia de Resultados, el cual fue presidido por Peter Fonagy, del University College de Londres (Fonagy, et al., 2004). El objetivo de este grupo de referencia fue guiar las buenas prácticas en la medición de resultados de salud mental, abordando algunas de las principales barreras para su implementación en la práctica.

Australia fue uno de los países pioneros a nivel mundial en evaluación rutinaria de resultados en los servicios de salud mental (Slade, 2002). La implementación de un protocolo de evaluación surgió a raíz de la estrategia nacional de salud mental en 1992, la cual enfatizó la calidad de servicio y su efectividad. Ningún país tenía implementada la evaluación de resultados de salud a nivel nacional, por lo que no tenían referentes internacionales en los que basarse.

Durante este periodo, Australia se fijó en sus homólogos extranjeros, los cuales estaban desarrollando un proyecto similar de evaluación rutinaria de resultados. Es el caso de Reino Unido, donde el Profesor Wing del Royal College of Psychiatrists de Londres estaba desarrollando la escala HoNOS (Wing et al, 1998), la cual estaba ganando cada vez más adeptos como instrumento candidato para la evaluación de resultados.

Wing pudo asesorar sobre el uso de esta medida a partir de su implantación preliminar como medida rutinaria en el Reino Unido. Esta escala se probó en Victoria, en 1996, en cinco servicios de salud mental de la zona. El ensayo se diseñó para examinar la utilidad del instrumento como medida rutinaria de resultados y para ayudar a establecer la recopilación y el análisis de datos de resultados. El personal clínico, tras recibir una amplia formación, evaluó a 2.137 usuarios que acudían a estos servicios, y durante los siguientes tres meses realizó un seguimiento de aproximadamente la mitad de ellos. Los ensayos indicaron que la medición rutinaria de resultados era posible, e identificaron los puntos fuertes y débiles de la escala HoNOS como instrumento de medida (Trauer et al., 1999).

Actualmente, en Australia el análisis y el reporte de los datos de resultados rutinarios se han vuelto gradualmente más sofisticados. Se ha realizado un esfuerzo considerable en cuestiones como definir qué es “cambio clínicamente significativo” y cuál es la mejor manera de medir dicho cambio (Burgess et al., 2009a, 2009b).

En Inglaterra, Nueva Zelanda y Australia pueden encontrarse importantes iniciativas y esfuerzos para adoptar ROA en los servicios de salud mental (Gilbody et al. 2002a), y la utilización rutinaria de medidas de resultados en los servicios de salud mental es obligatoria; entre ellas, la escala HoNOS ocupa un lugar central (Dogmanas, 2014).

En Nueva Zelanda, al igual que en Australia, hay un protocolo de evaluación de resultados implementado en la red de salud mental de todo el país. El Ministerio de Salud neozelandés, en su plan 2012-2017 “Rising to the Challenge: The Mental Health and Addiction Service Development”, centra su atención en la medición de resultados y los indicadores clave de desempeño para ayudar a desarrollar una cultura de resultados en los servicios de salud mental y adicciones. La estrategia del Ministerio con la escala HoNOS consiste en garantizar el buen cumplimiento de la recopilación de datos antes de introducir otras medidas obligatorias. La tasa general de cumplimentación del HoNOS, a partir de 2016, es del 80% para los centros comunitarios y de pacientes hospitalizados combinadas, y del 80% para las admisiones y altas de pacientes hospitalizados (Smith, et al., 2023).

La información sobre los resultados se utiliza a nivel local, regional y nacional para evaluar la eficacia de los servicios. Nueva Zelanda ha implementado un marco de medición de los resultados basado en cinco esferas principales: información clínica y de adicciones, maorí, información autoevaluada y funcionalidad. Las medidas clínicas utilizadas son las diferentes versiones de la escala HoNOS (HoNOS Family of Measures), todas ellas obligatorias para su recogida en los servicios de salud mental de Nueva Zelanda (Te Pou, 2021). La calidad de los datos es fundamental para poder utilizar la información de los resultados de forma eficaz. A través del trabajo multidisciplinar, se garantiza una recopilación de datos coherente entre médicos, los equipos y los servicios individuales.

El uso principal que se hace de los datos es a nivel individual. Los resultados de buena calidad permiten discutir la información de resultados con los propios usuarios del servicio, los cuales participan en la atención recibida y su tratamiento, y puede permitir obtener feedback sobre la recuperación del mismo (Te Pou, 2021; Burgess 2017). Compartir los resultados con los usuarios del servicio forma parte de un plan de atención colaborativo, lo cual también puede mejorar las oportunidades de participación de los whānau (allegados al usuario, familiares).

El uso secundario de los datos de resultados es a nivel agregado. A este nivel, los datos de resultados de buena calidad pueden ayudarnos a comprender mejor los cambios en la salud, el bienestar y las circunstancias de las personas que acceden a los servicios de salud mental de un determinado país o comunidad. Ofrecen información sobre la planificación, las actividades de mejora del servicio y las iniciativas de evaluación comparativa, y proporcionan una visión general del rendimiento de la organización en esos indicadores a lo largo del tiempo.

En Europa, en el año 2003, el European Opinion Research Group (EORG) publicó el estado de la salud mental de la población europea. Los resultados ofrecidos en este primer Eurobarómetro recogen diferencias en las puntuaciones de incidencia de problemas de salud mental, en la energía y vitalidad declaradas, y en el apoyo social percibido entre los 17 países o regiones europeas participantes. Las variables sociodemográficas y otras variables de contexto asociadas a los problemas de salud mental son más o menos esperadas y en su mayoría están de acuerdo con lo que se conoce sobre la distribución de estas características. Tres años más tarde, en el año 2006, la comisión de las comunidades europeas publicó un libro verde en el que abordaban diferentes estrategias para mejorar la salud mental de la

población europea. Señalan como importantes cuatro aspectos en los que la UE podría centrarse:

- 1- Promover la salud mental, sentando las bases en la infancia y la adolescencia y promocionando la salud mental entre los trabajadores y las personas mayores.
- 2- Prevenir las enfermedades mentales, como la depresión, al ser uno de los problemas de salud más graves de la UE, reducir los trastornos derivados del consumo de sustancias psicoactivas como las drogas o el alcohol y prevención del suicidio.
- 3- Mejorar la calidad de vida de las personas afectadas por enfermedades o discapacidades psíquicas integrándolas en la sociedad mediante la desinstitucionalización de los servicios y salud mental, y ofreciendo la prestación de estos servicios en la atención primaria, centros locales y hospitales generales.
- 4- Desarrollar un sistema de información, investigación y conocimientos sobre salud mental para toda la UE.

La evaluación rutinaria en los servicios de salud mental ha ido en aumento en los últimos 20 años. Además de Reino Unido, existe una extensa publicación de estudios de otros países como Alemania (Andreas, et al., 2010; Pushner, 2015), Italia (Parabiaghi, et al., 2005; Lovaglio and Monzani, 2012), Suiza (Egger, et al., 2020; Hoffman et al., 2022) o Países Bajos (De Beurs, et al., 2018; Bruins, et al., 2021) que recogen las ventajas de la ROM en sus servicios de salud. Sin embargo, en Europa no encontramos una implementación de evaluación de resultados tan sólida como la australiana.

1.4.1.- ROM en Salud Mental en España, País Vasco, y Bizkaia

En España no hay establecida una evaluación de resultados a nivel nacional como es el caso de Australia o Reino Unido. En el desarrollo de la reforma psiquiátrica en la década de los 70 no existía una evaluación sistemática de los resultados del proceso para las y los pacientes y sus familiares ni para los y las profesionales y el conjunto del sistema sanitario. En lo referente a la evaluación, es marcada la ausencia de información fiable desde sus inicios en todo el territorio del Estado para realizar un correcto seguimiento del proceso de cambio en la psiquiatría en el Sistema Nacional de Salud (Ministerio de Sanidad y Consumo, 2007).

A finales del año 2004 se asume el proyecto de elaborar la Estrategia en Salud Mental para el conjunto del Sistema Nacional de Salud (SNS), donde se proponen seis líneas estratégicas: 1) Promoción de la salud mental de la población, 2) Prevención de la enfermedad

mental y erradicación del estigma asociado a las personas con trastorno mental; 3) Atención a los trastornos mentales; 4) Coordinación interinstitucional e intrainstitucional; 5) Formación del personal sanitario e Investigación en salud mental; y 6) Crear un sistema de información en salud mental (Ministerio de Sanidad y Política Social, 2010). Desde entonces, el Ministerio Nacional de Salud ha publicado tres planes estratégicos dirigidos a la Salud Mental (2006, 2009-2013 y el actual 2022-2026); además de un plan de acción en salud mental en los años 2022-2024.

El sexto objetivo de la estrategia aprobada en el año 2006, Sistemas de información y evaluación en salud mental, plantea la necesidad de desarrollar un sistema de indicadores para estructurar la medición en los diferentes ámbitos de actuación para poder asegurar la calidad de la información empleada (Norma UNE66175, 2003). Gracias a estos indicadores y mediciones periódicas, se podrá disponer de datos objetivos que sustenten el avance hacia unos buenos resultados con relación a sus grupos de interés (Ministerio de Sanidad y Consumo, 2007).

Tradicionalmente, los sistemas de información en el ámbito de la salud mental se han denominado Registros de Casos Psiquiátricos (RCP), surgidos en la década de los años ochenta durante la reforma psiquiátrica y estaban caracterizados por estar centrados en los pacientes, ser longitudinales y recoger los contactos de los pacientes con un conjunto de servicios de un área geográfica delimitada (Pons et al., 1992). Surgen a raíz de la necesidad planteada por algunas Comunidades Autónomas de contar con instrumentos que permitieran la evaluación de los objetivos de la reforma, tratando de superar los sistemas de información de carácter puramente administrativo y centrados en los servicios. Entre ellos se encuentran el del País Vasco, que se creó en 1983, y el de Asturias, que funciona desde 1986, y se pueden considerar como una herramienta para evaluar la calidad de la atención y planificación de servicios (Pedreira-Massa y Eguiagaray, 1996).

La información existente actualmente sobre salud mental (El Conjunto Mínimo Básico de Datos – CMBD) no recoge la actividad asistencial en cada Comunidad Autónoma y no se encuentra agregada a nivel nacional. El CMBD de cada paciente consiste en un conjunto de variables codificadas en el momento del alta, que proporcionan los datos demográficos, la institución y servicio que lo atienden y su proceso asistencial (enfermedades y procedimientos). El CMBD al alta hospitalaria es actualmente el principal sistema de recopilación de información sobre actividad y morbilidad atendida existente en España. La amplia implantación del CMBD con criterios compartidos posibilita la comparación de datos

en el ámbito autonómico, nacional e internacional (Ministerio de Sanidad y Consumo, 2007), pero la explotación y amplitud de información disponible es muy limitada.

En España, las Comunidades Autónomas tienen transferidas las competencias de gestión de salud y de servicios sociales. Es por esto que a nivel de Comunidades Autónomas existen diferencias en la información obtenida. En Asturias, La Rioja y el País Vasco se dispone de un registro de casos psiquiátricos con datos clínicos, sociodemográficos, de actividad asistencial e itinerario de los usuarios que cubre toda la Comunidad Autónoma. En estas Comunidades Autónomas existe un sistema de información único que integra todos los dispositivos asistenciales de la red pública de salud mental. Estos sistemas generan información rutinaria a nivel de indicadores de gestión e indicadores epidemiológicos, centrados en cada paciente (Ministerio de Sanidad y Consumo, 2007).

En el País Vasco la salud mental y la asistencia psiquiátrica se asientan en tres hitos significativos. El primero, el Plan de Asistencia Psiquiátrica y Salud Mental de Euskadi de 1983 que, bajo el concepto de la psiquiatría comunitaria, posibilitó la transición del modelo manicomial al modelo comunitario. El segundo, la integración en 1986 de todos los recursos de asistencia psiquiátrica y salud mental dentro de la estructura del Servicio Vasco de Salud-Osakidetza. El tercer hito fue la aprobación de la Ley 14/1986, de 25 de abril, que estableció el ámbito comunitario como preferente en la prestación de la asistencia psiquiátrica (Gobierno Vasco, 2023).

La Comunidad Autónoma del País Vasco (CAPV) fue una de las primeras comunidades en España tanto en establecer una red de centros de salud mental que abarcara todo el territorio como, en un periodo inicial, en la creación de recursos intermedios (centros y hospitales de día) (Chicharro Lezcano, 2001).

Dentro de la CAPV, está la Red de Salud Mental de Bizkaia (RSMB), anteriormente “Salud Mental Extrahospitalaria de Bizkaia”. La RSMB se considera una Organización de Servicios Integrada (OSI). Se crea el 22 de marzo de 2010 y supuso la integración de los servicios y dispositivos en una única red, aglutinando centros de salud mental, estructuras intermedias y equipos de tratamiento asertivo comunitario, unidades hospitalarias de subagudos y de rehabilitación. El modelo de atención comunitario supone la integración y normalización de los servicios asistenciales psiquiátricos en la red sanitaria general, el apoyo a la integración social y normalización de la población afectada y la coordinación con los recursos sociosanitarios (Red de Salud Mental de Bizkaia, 2019).

La RSMB cuenta con programas de atención especiales como el programa de adicciones, trastornos de alimentación y de primeros episodios psicóticos (Lehenak), además de una atención a las personas con Trastorno Mental Grave (TMG) en materia de rehabilitación psicosocial desde los años 90; y su labor integrada se lleva a cabo a través de los Equipos de Tratamiento Asertivo-Comunitario (Ministerio de Sanidad, Servicios Sociales e Igualdad, 2012).

En el informe del estado actual de la prevención de la salud mental publicado en el año 2012, se exponía que en la RSMB había registrados cerca de 6000 usuarios con TMG. Casi 1500 de estos siguen una trayectoria de cronicidad (15%) y, en especial, cerca de 600 destacan por su complejidad y multi-problemática requiriendo de un abordaje altamente específico e intensivo. En este contexto, surgió la necesidad de contar con programas específicos que se adecuen a todo el itinerario psiquiátrico que requiere el paciente. Como estrategia general se configura el Plan Individualizado de Rehabilitación (PIR) que se centra en todas las áreas de necesidad desde donde se articulan todas las intervenciones.

El PIR es un instrumento mixto de evaluación/intervención. Se trata de un documento/expediente que regula el conjunto de actuaciones e intervenciones enfocadas en conseguir los objetivos establecidos por el equipo terapéutico para un paciente dado tras la evaluación rutinaria de sus necesidades, capacidades y psicopatología. En él se recogen las expectativas del proceso y actúa como guía del mismo (Ministerio de Trabajo y Asuntos Sociales (IMSERSO), 2007). Para llevarlo a cabo es necesaria la evaluación del proceso de cronicidad por dos motivos: el primero, conocer el punto de partida para ajustar las intervenciones de forma adecuada, y el segundo, contrastar los cambios encontrados. La evaluación inicial debería ser realizada mediante un instrumento sencillo, consensuado y multidisciplinar. Debe evaluar todas las áreas de interés y poder detectar áreas problema para diseñar intervenciones. Debe ser un instrumento con validez demostrada y de uso extendido a efectos de comparabilidad. La escala de evaluación global elegida, por ser polivalente y por su agilidad en la cumplimentación rutinaria, es la escala HoNOS, traducida al castellano por Uriarte et al. (1999).

Capítulo 2

HoNOS - Health of the Nation Outcome Scale

La utilización de instrumentos psicométricos en la evaluación de resultados en salud mental es una opción que ha venido utilizándose durante décadas. Existe todo un repertorio de instrumentos y escalas de medida (Slade, 2002; McDowell, 2006) orientados a la valoración de signos y síntomas psíquicos específicos (depresión –BDI, HDS– o alexitimia –TAS-20–, por ejemplo) o a la valoración de una generalidad de manifestaciones valorados a través de instrumentos multicomponentes (SCL-90, GHQ, OQ-45, CORE-ON y/o SF36). Para su correcta utilización en el ámbito de evaluación de resultados resulta imprescindible que dichos instrumentos se hayan adaptado adecuadamente al idioma y cultura en la que se utilicen (Wild et al., 2009; Morkink et al., 2010) y sea probada su capacidad y adecuación de medida en los términos clásicos de validez y fiabilidad (Scholtes et al., 2001).

Uno de los instrumentos más utilizados a nivel internacional, y más versátiles en su aplicación clínica en el trastorno mental grave, es el HoNOS (Wing et al., 1998), al cual nos aproximaremos en este capítulo.

2.1.- La escala HoNOS

2.1.1.- Génesis y Desarrollo

La escala HoNOS (Health of the Nation Outcome Scale) fue desarrollada a principios de la década de los 90 en Gran Bretaña por la Unidad de Investigación de The Royal College of Psychiatrists' Research Unit en el marco del Proyecto denominado “Health of the Nation Strategy”, del Departamento de Salud (DS) del Reino Unido (Department of Health, 1992; Wing, et al., 1998). Se identificaron cinco áreas clave para el desarrollo de estrategias

orientadas a reducir la mortalidad y morbilidad en Reino Unido. Estas áreas seleccionadas fueron: cardiopatías y accidentes cerebrovasculares, cáncer, salud sexual, los accidentes y la salud mental (Trauer, 1998). La escala se desarrolló con el propósito de constituirse en una medida de resultados sensible, con una amplia cobertura de problemas que pudiera ser utilizada de manera rutinaria en los servicios de salud mental de adultos para cuantificar la mejoría de la salud y el funcionamiento social de las personas con enfermedades mentales (Pirkis, et al., 2005; Wing et al., 1998). Actualmente, el área de la Salud Mental sigue siendo considerada como un área sanitaria de importancia prioritaria (Burgess, et al., 2012).

En el documento *Health of the Nation* se desarrollan primero dos objetivos en relación con la salud mental: reducir los índices de suicidio tanto en la población general, así como en la población clínica. Sin embargo, como el índice de suicidio no es un buen indicador de la salud mental global de un país, decidieron añadir un tercer objetivo, el de “mejorar significativamente la salud y el funcionamiento de las personas con enfermedades mentales” (Department of Health, 1992). Este objetivo puede expresarse en términos de mejoría clínica y/o social, el cual se acerca al concepto de “autonomía funcional óptima” por el mantenimiento de un estado de funcionamiento óptimo, previniendo, frenando y/o mitigando el deterioro. De este modo, para establecer un objetivo cuantificado y medir su grado de cumplimiento, se hacía necesario un instrumento específico que pudiera utilizarse de forma rutinaria en el Servicio Nacional de Salud (SNS) (Trauer, 1998).

El Departamento de Salud buscaba una recogida de datos rápida, sencilla de utilizar y de puntuar, aplicable por cualquier profesional de salud mental de forma rutinaria en cualquier tipo de dispositivo asistencial. El instrumento por desarrollar debía poder utilizarse para combinar los datos recogidos y evaluar si la salud física, social y mental del sujeto mejoraba a lo largo del tiempo. Otro de los requisitos era que el instrumento permitiera realizar comparaciones entre diferentes servicios y que pudiera servir a los proveedores de servicios para ver cómo los diferentes tipos de unidades afectaban a la salud mental y social de los pacientes (Wing et al., 1998; Trauer, et al., 1998).

2.1.2.- Descripción y características de aplicación

La escala HoNOS consiste en una medida basada en el juicio del clínico (*clinician rated*). Está formada por 12 ítems que cubren una gama de problemas biológicos, psicológicos y sociales (ver más adelante Tabla 2.1, para apreciar la especificidad de los ítems). Los autores de la escala (Wing, 1998) proponen agrupar los ítems en cuatro

subescalas: problemas conductuales (ítems 1, 2 y 3); deterioro (4 y 5); problemas clínicos (6, 7 y 8) y problemas sociales (9, 10, 11 y 12). La severidad o gravedad con la que se expresa cada ítem se mide en una escala de cinco opciones de respuesta, de 0 a 4, donde una mayor puntuación correspondería a una mayor gravedad [0 - No hay problema; 1 - Problema menor que no requiere ninguna acción; 2 - Problema leve pero definitivamente presente; 3 - Problema moderadamente grave; 4 - Problema grave a muy grave]. El rango de puntuaciones posibles, por tanto, oscilaría entre cero (nula expresión sintomática) a 48 (en el caso de que la persona evaluada expresara a máxima gravedad en cada una de las 12 áreas exploradas). Para la correcta valoración de la intensidad atribuible existen un programa de entrenamiento y manuales de apoyo que sustentan la formación en la aplicación del instrumento (Uriarte et al., 1999; Te Pou, 2021).

Según Delaffon, Anwar, Noushad, Ahmed, y Brugha (2012) el tiempo requerido para completar la escala es de 2-3 minutos. Algunos autores reportan tiempos de empleo mayores, de 5 a 15 minutos (McClelland, 2000), y de entre 15-30 minutos (Pirkis, et al., 2005). Estas variaciones en el tiempo de aplicación se basan en las diferencias entre los clínicos, la familiaridad que tengan con la escala y la disponibilidad de la información necesaria (Delaffon et al., 2012).

La HoNOS es completada por un profesional de la salud mental cualificado (clínico) utilizando la información obtenida tanto en la evaluación de la salud mental como de su trabajo clínico diario. Se recomienda tener en cuenta la información de todas las fuentes disponibles al completar las calificaciones, incluida la información proporcionada por el evaluado, sus familiares o personas cercanas y también notas clínicas recogidas (Te Pou, 2021).

La HoNOS se califica utilizando un extenso glosario con instrucciones detalladas para realizar la evaluación (Anexo 1) y se basa en la información de las dos últimas semanas de presentación de un usuario del servicio en todas las ocasiones de recogida. Como excepciones, al finalizar un episodio en un entorno de hospitalización el período de calificación es de tres días; a su vez, las escalas (ítems) 11 y 12 pueden calificarse más allá del período de dos semanas (Te Pou, 2021).

La escala HoNOS se completa en tres situaciones: a) cuando una persona entra en un servicio especializado de salud mental (al ingreso o al inicio de un episodio) y cuando sale (alta o fin del episodio) del servicio; b) en tres periodos de revisión mensuales mientras sigan

accediendo a los servicios; y c) cuando se produzcan cambios significativos en la salud, el bienestar o las circunstancias del usuario del servicio (revisión ad hoc) (Te Pou, 2021).

Un clínico referente, que es el que mejor conoce al usuario del servicio, registra las calificaciones de la HoNOS, teniendo en cuenta toda la información disponible. Lo ideal es que el mismo clínico o equipo califique la revisión posterior o de alta de la HoNOS. Esto no siempre es posible, sobre todo en el caso de los centros de tratamiento hospitalario (Te Pou, 2021). En el caso de que la escala no sea aplicada por el clínico que realiza el tratamiento, previamente es necesario tener una entrevista semi-estructurada (Dogmanas, 2014).

Si el usuario es admitido en una unidad de hospitalización cuando se encuentra en los servicios comunitarios, se requerirá la recopilación de datos de admisión y alta para la hospitalización, pero el episodio comunitario no se cerrará. Este se convierte en el escenario principal y permanece abierto si hay un ingreso hospitalario. Por lo tanto, la fecha de revisión será fijada por el ingreso comunitario y no por el ingreso de hospitalización. Si hubiera que hacer tres evaluaciones mensuales de revisión mientras el paciente está hospitalizado, sería deseable que miembros sanitarios del equipo comunitario se encargaran de realizarlas (Te Pou, 2021).

2.1.3.- Estructura Dimensional del HoNOS

Como se ha comentado en el apartado anterior, en su propuesta final original (Wing et al., 1998), los autores proponen un instrumento de 12 ítems que explorarían cuatro dimensiones en las que se manifiestan los trastornos mentales graves: problemas conductuales, deterioro, problemas clínicos y problemas sociales. La investigación posterior desarrollada por otros autores sobre el HoNOS ha mostrado otras alternativas de clasificación de los ítems (Tabla 2.1) que, respondiendo a su configuración matemática obtenida mediante técnicas de análisis factorial, ofrece una amalgama de posibles configuraciones conceptuales de las manifestaciones de resultados en trastorno mental grave: Trauer (1999) propone el uso de cinco factores, otros autores (Preston, 2000; McClelland et al., 2000; Newnham et al., 2009; Speak & Muncer, 2016) encuentran cuatro factores correlacionados distintos al original, Lovaglio y Monzani (2012) un único factor con seis ítems y Speak y Muncer (2015) un factor con cuatro ítems que hacen referencia al factor social del modelo original. En la base de todas estas discrepancias se halla el propio comportamiento de los ítems que ofrecen una correlación moderada- baja entre los mismos (Trauer, 1999).

Tabla 2.1.- Estructuras factoriales sugeridas para el HoNOS

Wing (1998)	Ítems	Trauer (1999)	Preston (2000)	McClelland (2000)	Newnhan (2009)	Lovaglio (2012)	Speak (2015)	Speak (2016)
F1 –	It01 – Agresividad	1	1	1	1			4
Problemas conductuales	It02 – Autoagresiones	3	1	2	1			2
	It03 – Uso de sustancias	1	1	1	1			3
F2 –	It04 – Disfunción cognitiva	2	2	1	2	1		1
Deterioro	It05 – Discapacidad Física	2	2	3	2			1
F3 –	It06 – Depresión	3	3	1	3	1		4
Problemas Clínicos	It07 – Alucinaciones/Delirios	5	3	2	2			2
	It08 – Otros Síntomas	3	3	4	3			2
F4 –	It09 – Relaciones Sociales	3-4	4	1	1	1	1	3
Problemas sociales	It10 – Funcionamiento general	4	4	1	2	1	1	1
	It11 – Problemas de residencia	4	4	1	4	1	1	3
	It12 – Problemas ocupacionales	4	4	1	4	1	1	1-3

Nota. Los factores de Preston, McClelland mantienen las denominaciones propuestas por Wing en el estudio original. Trauer define los factores de la siguiente forma: F1=Conducta, F2=deterioro, F3=Depresión, F4=Problemas sociales y F5=Alucinaciones. Newnhan los define: F1=Problemas conductuales o antisociales, F2=Deterioro, F3=Síntomas y F4=Socioeconómico. Speak (2016) los define: F1=Bienestar personal, F2=Bienestar emocional, F3=Bienestar social y F4=Alteraciones severas.

En su elección inicial (Wing et al., 1998) se optó por una escala breve que reflejase todo el abanico de manifestaciones en el ámbito del trastorno mental grave, imponiéndose las consideraciones prácticas de medir el fenómeno sobre las matemáticas. No obstante, en los últimos años se ha venido promoviendo una alternativa en la valoración de la configuración de medida de los instrumentos basada en el análisis bifactorial (Reise et al., 2013; Rodríguez & Reise, 2016) que pudiera ser una alternativa en la comprensión del comportamiento psicométrico del HoNOS.

El uso del HoNOS en España se ha posibilitado a través de la adaptación realizada por el grupo de Uriarte en colaboración con los autores del instrumento (Uriarte et al., 1999). Si bien ha sido utilizado en distintos estudios (Gisbert & Cid- Colom, 2010; Frades et al., 2011; Osta-Samanes et al., 2012; Garay-Arostegui et al., 2014; Escalada et al., 2014, 2015a, 2015b; Ribas et al., 2016) y a excepción del estudio de validación en población uruguaya (Dogmanas,

2014) y de una adaptación realizada para personas con discapacidad de aprendizaje (Esteban-Castillo, et al., 2018), a día de hoy no se conoce ningún estudio que informe sobre sus características de medida y adecuación psicométrica al español, siendo éste uno de los objetivos de la presente tesis.

2.1.4.- Versiones del HoNOS

La escala HoNOS original se desarrolló para los servicios de salud mental de adultos que se encuentran en edad laboral, de 18 a 64 años, ambos incluidos. Posteriormente, se han desarrollado versiones para otras edades: HoNOSCA (HoNOS for Children and Adolescents, versión para niños y adolescentes); HoNOS +65 (versión para adultos mayores de 65 años) y HoNOS-OA (HoNOS Older Adults, versión revisada en 2018 de la escala original HoNOS +65).

Asimismo, existen versiones para otros servicios específicos, que se repasan a continuación.

HoNOS-secure (para servicios forenses) (Dickens, et al., 2007) versión revisada de la anterior versión HoNOS-MDO (Mentally Disordered Offenders). La segunda versión de HoNOS-Secure se desarrolló a finales de 2004 simplificando las instrucciones originales y realizando algunos cambios en la redacción de las escalas de seguridad, obteniendo en 2007 la versión final revisada. A los 12 ítems de la escala original se suman las siete escalas de valoración de la seguridad: A = riesgo de daño a adultos o niños, B = riesgo de autolesión (deliberada o accidental), C = necesidad de seguridad del edificio para evitar la fuga física, D = necesidad de un entorno vital con personal seguro, E = necesidad de escolta en la salida (más allá del perímetro de seguridad), F = riesgo para el individuo por parte de otros, G = necesidad de procedimientos de gestión de riesgos. Al igual que las versiones anteriores, la versión 2b de HoNOS-secure no es una herramienta de evaluación de riesgos. Permite calificar el resultado de la evaluación del riesgo clínico en términos de necesidad de cuidados y de necesidad de procedimientos de gestión del riesgo clínico (Sugarman and Walker, 2007).

HoNOS-LD para personas con problemas de aprendizaje (Roy, Matthews, Clifford, Fowler, y Martin, 2002). Esta escala está diseñada para su uso con personas con discapacidad de aprendizaje con necesidades de salud mental, independientemente del grado de su discapacidad. La HoNOS-LD ha sido sometida a pruebas de fiabilidad entre evaluadores, fiabilidad convergente, sensibilidad al cambio y aceptabilidad en un estudio piloto nacional.

Los resultados demuestran una buena fiabilidad entre evaluadores y un nivel de sensibilidad al cambio aceptable para los clínicos que trabajan en este campo.

HoNOS-ABI (HoNOS-Acquired Brain Injury), la escala para personas con daño cerebral adquirido (Fleminger et al., 2005) fue derivada en gran medida de HoNOS+65 por un grupo de psiquiatras especializados en lesiones cerebrales interesados en desarrollar una versión HoNOS más adecuada para este grupo de personas.

HoNOS-RF (Residential Facilities) el equipo de Pedrini et al, (2017) desarrolló y validó esta versión con el fin de llevar a cabo una evaluación exhaustiva de los pacientes ingresados en residencias psiquiátricas. Consta de 31 ítems, agrupados en ocho factores: funcionamiento personal e interpersonal, entorno, comportamiento y carga de cuidados, función cognitiva, problemas somáticos, síntomas ansioso-depresivos, síntomas psicóticos y otros síntomas psiquiátricos.

Otras modificaciones de la HoNOS se produjeron en una versión de autoinforme en el marco del paradigma de recuperación (Stewart, 2009). La versión HoNOS-MHCT (Mental Health Clustering Tool) y la versión Honos Payment by Results (HoNOS-PbR). El MHCT se ha desarrollado en colaboración con el Departamento de Salud Británico, el Royal College of Psychiatrists Centre for Advanced Learning and Conferences y el Care Pathways and Packages Project (CPPP) como medio para asignar a los clientes a grupos de atención, lo que a su vez apoya la planificación de la atención y permite el Pago por Resultados en Salud Mental (MH PbR-Payment by Results). El MHCT mantiene la escala HoNOS y, por lo tanto, ayuda a los proveedores y a los comisionados a medir los resultados de la atención sanitaria y social en materia de salud mental sin necesidad de recopilar datos adicionales.

2.2.- Uso internacional y vinculación a servicios de salud

La incorporación del uso de HoNOS como medida de resultado clínico rutinario aumenta de manera progresiva (Parabiaghi, 2014), lo cual permite, además, monitorear, juzgar y mejorar la calidad y efectividad de los servicios de salud mental (Burguess, 2012). La escala forma parte del conjunto mínimo de datos nacionales que deben recolectar en los procesos de admisión, seguimiento y alta terapéutica en los servicios de salud mental británicos, en Australia, Nueva Zelanda y Países Bajos, por lo cual su uso es obligatorio (Pirkis et al, 2005; Parabiaghi, 2014; Burgess, 2012; De Beurs, 2011). También es un instrumento de referencia en diferentes regiones de Canadá, Dinamarca, Francia, Italia, Alemania, Holanda y Noruega

(Andreas et al., 2011; Pirkis, Burgess, Kirk, et al., 2005) además de haber sido adaptada y validada en 15 países (Iglesias et al, 2019).

El uso de la HoNOS está recomendado por el Marco Nacional de Servicios de Salud Mental Inglés y por el grupo de trabajo del Departamento de Salud sobre indicadores de resultados para enfermedades mentales graves (Jacobs, 2009). La escala HoNOS ha sido diseñada para utilizarse con pacientes con trastornos mentales de los servicios de salud mental de segundo y tercer nivel, con una amplia gama de diagnósticos (Andreas et al., 2011) y en los servicios de atención ambulatorio, de internación psiquiátrica y residencial. Se puede aplicar también en la evaluación en atención primaria, aunque su uso con pacientes menos severos está menos estudiado (Pirkis, Burgess, Kirk, et al., 2005). Por otra parte, Audin, Margison, Clark y Barkham (2001) reportaron poca utilidad en evaluar a pacientes atendidos en los servicios de psicoterapia que presentan patologías menos severas.

La HoNOS es la medida de resultados clínicos de rutina más utilizada por los servicios de salud mental ingleses. La Royal College of Psychiatrists' Research Unit (CRU) realizó una encuesta en los servicios de salud mental del NHS en Inglaterra en el año 2000 y encontró que 34 servicios habían implementado el uso de HoNOS en uno o más entornos de servicio. Otra encuesta realizada por el Departamento de Salud en 2002 descubrió que 61 empresas utilizaban la escala HoNOS en sus servicios y que cinco habían implementado el uso de las escalas en todo el servicio (Jacobs, 2009).

Está considerado como uno de los instrumentos más utilizados a nivel internacional, y forma parte del conjunto de datos básicos de recolección obligatoria en los países que poseen mayor desarrollo en el campo de la evaluación de resultados en servicios de salud mental. Hasta el momento, la escala HoNOS ha sido traducida a nueve idiomas, entre ellos alemán, francés, checo, danés, italiano y tailandés (Delaffon et al., 2012; Pirkis, Burgess, Kirk, et al., 2005) y cuenta con numerosos estudios de validación realizados en 15 países (Iglesias et al, 2019). El equipo de Uriarte et al. (1999) realizó la traducción oficial al castellano en 1999 junto con los autores originales de la escala. Recientemente se ha publicado el estudio de validación de la escala en población española (Uriarte et al., 2024), que forma parte de esta tesis (Estudio 1).

2.2.1.- Resultados a nivel internacional

Desde el desarrollo de la escala, se han realizado diferentes estudios sobre las propiedades psicométricas de la escala HoNOS hasta en un total de 15 países (Iglesias et al,

2019). Los países que cuentan con estudios de validación de HoNOS son Reino Unido (Wing, 1998; Orrell, 1999; Richardson, 2019), Australia (Goldney, 1998; Trauer, 1999; Parker, 2020), Nueva Zelanda (Eagar, 2005; Jury, 2018), España (Uriarte, 2020), Italia (Parabiaghi, 2005; Lovaglio, 2012), Países Bajos (Mulder, 2004; Kwarkenaak, 2019), Dinamarca (Bech, 2003), Noruega (Øiesvold, 2011), Alemania (Andreas, 2007), Suiza (Theodoridou, 2009; Böckmann, 2019), Canadá (Lauzon, 2001; Currie, 2020), Irlanda (Browne, 2000; Sweeney, 2019), Republica Checa (Pec, 2009), Indonesia (Anjara, 2019) y Uruguay (Dogmanas, 2014).

Si bien los siguientes países no presentan estudios sobre las propiedades psicométricas de la escala, sí se ha usado la HoNOS para estudios clínicos en USA (Muehsam, 2018), Bélgica (Lorant, 2018), Hong Kong (Hariman, 2020) y Grecia (Peritogiannis, 2018).

Se han realizado diversas revisiones sistemáticas que recogen datos del uso de la escala HoNOS (Agius, 2010; Delaffon, et al., 2012; Burgess, Harris, Coombs, et al., 2017; Harris et al, 2018; Iglesias-López et al, 2019) o sus propiedades psicométricas (Pirkis, Burgess, Kirk, et al., 2005; Delaffon et al, 2012; Harris et al, 2018) así como revisiones de la función que desempeña el instrumento dentro de los servicios de salud mental en un país, por ejemplo Australia (Callaly, Trauer and Hantz, 1998; Pirkis, Burgess, Coombs, et al., 2005) o Reino Unido (Mental Health Bulletin, 2014).

La revisión que realiza Pirkis, Burgess, Kirk, et al., (2005) recoge estudios que presentan evidencias sobre la fiabilidad de la escala (alfa de Cronbach entre .59 y .76) lo cual indicaría una consistencia interna moderadamente alta, confirmando su uso como instrumento de medida de severidad en salud mental (Wing et al., 1998; Shergill, 1999; Eagar, et al., 2005). En la misma revisión, los autores señalan que la HoNOS demuestra ser sensible al cambio, reportando un mayor cambio en pacientes internados respecto a los ambulatorios, así como en pacientes que presentan puntuaciones totales mayores al inicio del tratamiento (Pirkis, Burgess, Kirk, et al., 2005; Delaffon et al., 2012).

En referencia a la validez convergente, existen numerosos estudios que presentan correlaciones positivas entre la escala HoNOS y otros instrumentos de medida, como son the Role Functioning Scale (RFS); Brief Psychiatric Rating Scale (BPRS); Life Skills Profile (LSP) (Pirkis, Burgess, Kirk, et al., 2005) y el CANSAS (Camberwell Assessment of Need Short Appraisal Schedule), SCAN (Schedules for Clinical Assessment in Neuropsychiatry), SBS (Social Behaviour Scale) y el SCL-90-R (Symptom Checklist 90 Revised) (Iglesias-López, 2019).

Si bien es apreciable el número considerable de países que han validado el instrumento o lo están utilizando en la práctica asistencial, una revisión sistemática realizada en 2012 por el equipo de Delaffon, sugería la necesidad de realizar más estudios psicométricos de la escala HoNOS ante la presencia de datos no convergentes (Delaffon et al., 2012).

Capítulo 3

Revisión sistemática de las evidencias de validez y fiabilidad de la escala HoNOS: adaptaciones y validaciones internacionales de la escala

En conjunto, los múltiples estudios realizados con el HoNOS han mostrado buenas propiedades psicométricas (Pirkis et al., 2005; Delaffon et al, 2012; Harris et al, 2018), lo que ha afianzado la decisión de incorporarlo como medida de resultado clínico rutinario (Parabiaghi, 2014) permitiendo, además, monitorear, juzgar y mejorar la calidad y efectividad de los servicios de salud mental (Burguess, 2012).

A pesar de existir revisiones sistemáticas previas sobre el HoNOS (Pirkis et al 2005; Delaffon, 2012; Agius, 2010; Ansvarlig Magne Nylenna, 2014; Burgess, 2017; James, 2018; Harris, 2018), ningún estudio presenta un análisis completo de bondad psicométrica de la escala HoNOS, y en muchos de los estudios clínicos se omite información sobre la fiabilidad alcanzada por la escala. A este respecto, como ya se ha comentado previamente, la revisión sistemática del equipo de Delaffon (2012), sugería la necesidad de realizar más estudios psicométricos de la escala HoNOS ante la presencia de datos no convergentes.

El objetivo de este capítulo, por tanto, es realizar una revisión sistemática actualizada de las evidencias de comportamiento psicométrico del HoNOS a nivel internacional siguiendo las recomendaciones del protocolo PRISMA-P (Shamseer, et al., 2015) y tratando de ofrecer una integración meta-analítica de de las evidencias de su capacidad como medida de resultados en salud mental.

3.1.- Método

3.1.1.- Diseño

Se trata de un estudio de revisión sistemática con integración meta-analítica de datos, en el que se siguieron las recomendaciones del protocolo PRISMA-P (Shamseer, et al., 2015) y los cálculos de integración meta-analítica realizados con JAMOVI (The Jamovi Project, 2020). Este estudio se registró en PROSPERO con el código CRD42020171115.

3.1.1.1.- Estrategia de búsqueda

La revisión sistemática se llevó a cabo usando las bases de datos bibliográficas EBSCOHost, PsicINFO, PsicARTICLES, PubMed, Science Direct, Web of Science, SCOPUS, Elsevier y Embase entre los años 1998 y 2023. La ecuación de búsqueda se formó con el nombre completo de la escala *Health of the Nation Outcome Scale*, sus siglas *HoNOS* y los términos *Psychometric Properties*, *Validity*, *Reliability*, *Adaptation* y *Construct Validity*. La última búsqueda en las bases de datos se realizó en octubre de 2023. Además, las referencias de los estudios seleccionados, así como anteriores revisiones del instrumento fueron analizadas para identificar otras posibles publicaciones relevantes.

3.1.1.2.- Criterios de inclusión

- a) Los estudios debían emplear la última versión para adultos HoNOS-4.
- b) Debían aportar datos psicométricos de fiabilidad y/o validez del instrumento, y/o datos clínicos que cuantificaran la gravedad de la muestra (Media y Desviación típica).
- c) La muestra del estudio debía estar formada por población clínica.

3.1.1.3.- Criterios de exclusión

- a) Emplear otras versiones del instrumento: HoNOSCA, HoNOS +65, HoNOS-RF, HoNOS-LD, HoNOS-ABI y HoNOS-Secure.
- b) Estudios que, aun incluyendo una versión para adultos, esta difería de la original de Wing et al (1998)⁵: HoNOS-3 o adaptaciones del HoNOS-4 (HoNOS-Roma y HoNOS-Thai). La adaptación de la escala HoNOS en estos países concluyó en una alteración del número de ítems respecto a la versión original (18 ítems y 10 ítems respectivamente).

- c) Estudios de revisión.
- d) No aportan datos psicométricos consistentes o concluyentes que dificultan su integración.

3.1.1.4.- Proceso de selección y extracción de datos

Un investigador se encargó de revisar los títulos y resúmenes de las publicaciones con el fin de excluir aquellos estudios que no cumplían con los criterios de inclusión/exclusión, sin restricción del idioma. La búsqueda de estudios se realizó en cuatro tiempos. La primera revisión se realizó en octubre de 2018, y la búsqueda se actualizó en octubre de 2019, abril de 2020 y octubre de 2023. Una vez revisados, se obtuvieron los artículos de los estudios seleccionados y dos revisores de forma independiente analizaron el texto completo en base a los criterios anteriormente descritos, con el fin de contemplar su relevancia. Las discrepancias se resolvieron mediante debate o, cuando no se pudo llegar a un consenso, recurriendo a un tercer examinador.

3.1.2.- Calidad metodológica

Las cualidades métricas cuantitativas de los estudios seleccionados fueron evaluadas de acuerdo con los criterios de adecuación de instrumentos de medida de salud predefinidos en el listado COSMIN “checklist” (COnsensus-based Standards for the selection of health status Measurement Instruments) (Mokkink et al, 2010, 2010b, 2010c, 2018; Prinsen et al, 2018, Terwee et al., 2007, 2012, 2018). En base a las recomendaciones de COSMIN, se realizaron síntesis de las evidencias encontradas para cada propiedad de medición, teniendo en cuenta las calificaciones de calidad metodológica y los datos de cada una.

Para la aproximación a los criterios que definen las propiedades psicométricas de un instrumento de medida, la taxonomía de COSMIN (Mokkink et al. 2010c) está constituida por tres dominios generales: fiabilidad, validez y receptividad (responsiveness), en los cuales se incluyen las propiedades estadísticas de medida.

3.1.2.1.- Fiabilidad

Hace referencia al grado en que la medición se encuentra libre de error de medida. Dentro de esta categoría encontramos la fiabilidad (versión ampliada), la consistencia interna y el error de medida.

Fiabilidad - Es la medida en que las puntuaciones de los pacientes que no han cambiado permanecen estables en repetidas evaluaciones en condiciones diferentes. Por ejemplo, aplicando diferentes grupos de ítems de un mismo instrumento (consistencia interna), aplicando el instrumento en dos momentos diferentes y seriales (test-retest), por diferentes personas en la misma situación (inter-evaluador, por ejemplo, a través el índice de Kappa) o por las mismas personas en diferentes momentos (intra-evaluador).

Consistencia interna - Es definida como el grado en el que los ítems se intercorrelacionan entre sí, midiendo de esta manera el mismo constructo (Mokkink, et al, 2010). El panel de expertos del estudio Delphi concluyó que la consistencia interna solo alcanza un significado interpretable en dos ocasiones: la primera cuando la intercorrelación entre los ítems está determinada por el conjunto de ítems formando un modelo reflectivo, y la segunda cuando todos los ítems intervienen en el mismo constructo, formando por ejemplo una escala unidimensional.

Error de medida - Es el error sistemático y aleatorizado en las puntuaciones de un paciente que no son atribuidas a cambios reales en el constructo que se quiere medir. En este estudio calculamos el error estándar de medida (SE) a través de la siguiente fórmula:

$$SE = \sigma / \sqrt{n};$$

siendo σ la variabilidad de la distribución de la medida y n el tamaño muestral.

3.1.2.2.- Validez

Definida como el grado en que un instrumento de evaluación de salud mide el constructo que pretende medir, presenta diferentes acepciones.

Validez de contenido – Es el grado en el que el contenido de un instrumento refleja de forma adecuada el constructo que pretende medir. De acuerdo con el panel de expertos (Mokkink, et al, 2010b), la validez de contenido debe medirse por la relevancia de los ítems y su capacidad de ser comprendidos. Dentro de la validez de contenido se encuentra la *validez aparente*, la cual comprende el grado en el que un instrumento aparenta ser un adecuado reflejo del constructo que pretende medir (Terwee, 2007)

Validez de constructo – Es el grado en que las puntuaciones de un instrumento de evaluación de salud son consistentes con las hipótesis teóricas basadas en el supuesto de que el instrumento mide de forma válida el constructo que se quiere medir. Dentro de la validez de constructo encontramos tres tipos de medidas psicométricas. La *validez estructural* o

factorial, la cual consiste en el grado en el que las puntuaciones de un instrumento de evaluación de salud son un adecuado reflejo de la dimensionalidad del constructo que pretende ser medido. Segundo, *la prueba de hipótesis*, la cual refleja la propia validez de constructo. Por último, la *validez intercultural* es el grado de rendimiento en el que los ítems de un instrumento traducidos o adaptados culturalmente son un adecuado reflejo de la ejecución de los ítems en la versión original del mismo instrumento.

Validez de criterio - La validez de criterio es el grado en el que las puntuaciones de un instrumento son un adecuado reflejo del criterio de referencia (gold standard).

3.1.2.3.- Capacidad de respuesta (*responsiveness*)/Sensibilidad al cambio

La capacidad de respuesta es la habilidad de un instrumento de evaluación de detectar cambios en el tiempo en el mismo constructo que se mide.

Interpretabilidad - Es el grado por el cual una persona puede atribuir un significado cualitativo clínico a un instrumento con puntuaciones cuantitativas o cambios en sus puntuaciones.

3.1.3.- Integración meta-analítica de datos clínicos

Se han realizado meta-análisis utilizando programas de integración (plantillas Excel) para generar estimaciones conjuntas de las mediciones de las propiedades psicométricas pertinentes (*r* de Pearson-validez concurrente, Alpha de Cronbach-fiabilidad, y SEM-error sistemático de medida). Estos datos se han integrado, por una parte, con los estudios de cada país y, por otro, el conjunto total de estudios. Teniendo como referencia revisiones sistemáticas previas, se han analizado las conclusiones referentes a los indicadores de bondad psicométrica cualitativa; éstos se han sintetizado de forma narrativa (Anexos 3 y 4).

Se llevó a cabo una integración de resultados de los datos de gravedad clínica disponibles extraídos de los estudios (medias ponderadas, desviaciones estándar -DE- y el error estándar de medida -EEM-) mediante el cálculo ponderado a través de una aplicación en hoja Excel. Los coeficientes de fiabilidad -Alpha de Cronbach- se integraron a través del software estadístico JAMOV.

Los puntos de corte (PC) de la escala y el cambio mínimo significativo (CMS) se calcularon a partir de las fórmulas de Jacobson y Truax (1991) en su propuesta de Índice de Cambio Fiable (RCI). Este índice permite calcular los efectos que tiene un tratamiento a nivel

individual, con la proporción de cambios fiables y clínicamente significativos observados con una evaluación pre-tratamiento y otra post-tratamiento (ej. ingreso y alta médica) (Maxwell et al, 2018). En primer lugar, se calculó el índice fiable de cambio para determinar si la puntuación de cambio de cada paciente individual era fiable, o más allá de lo que podría esperarse que ocurriera por azar (Jacobson y Truax, 1991), utilizando la fórmula:

$$RCI = \frac{Puntuación\ PostTratamiento - Puntuación\ PreTratamiento}{SEdiff}$$

donde

$$SEdiff = \sqrt{2(SD_1\sqrt{1-\alpha})^2},$$

siendo SD1 la desviación estándar de la puntuación pre-tratamiento y α el coeficiente de Cronbach de fiabilidad interna de la medida de resultado pre-tratamiento. Un índice de cambio fiable puede interpretarse estadísticamente como una puntuación z estándar, donde valores superiores a 1,96 que indican que se ha producido un cambio con baja probabilidad de ser debido al azar en una hipótesis de dos colas (Duff, 2012).

La significación clínica hace referencia a la medida en que las puntuaciones posteriores al tratamiento han cruzado un umbral o puntuación mínima para ser clasificadas como más cercanas a la distribución del grupo de comparación de la población normativa o funcional que a la media previa al tratamiento de la muestra clínica. Para ello es necesario determinar un punto de corte (PC), calculado con la siguiente fórmula:

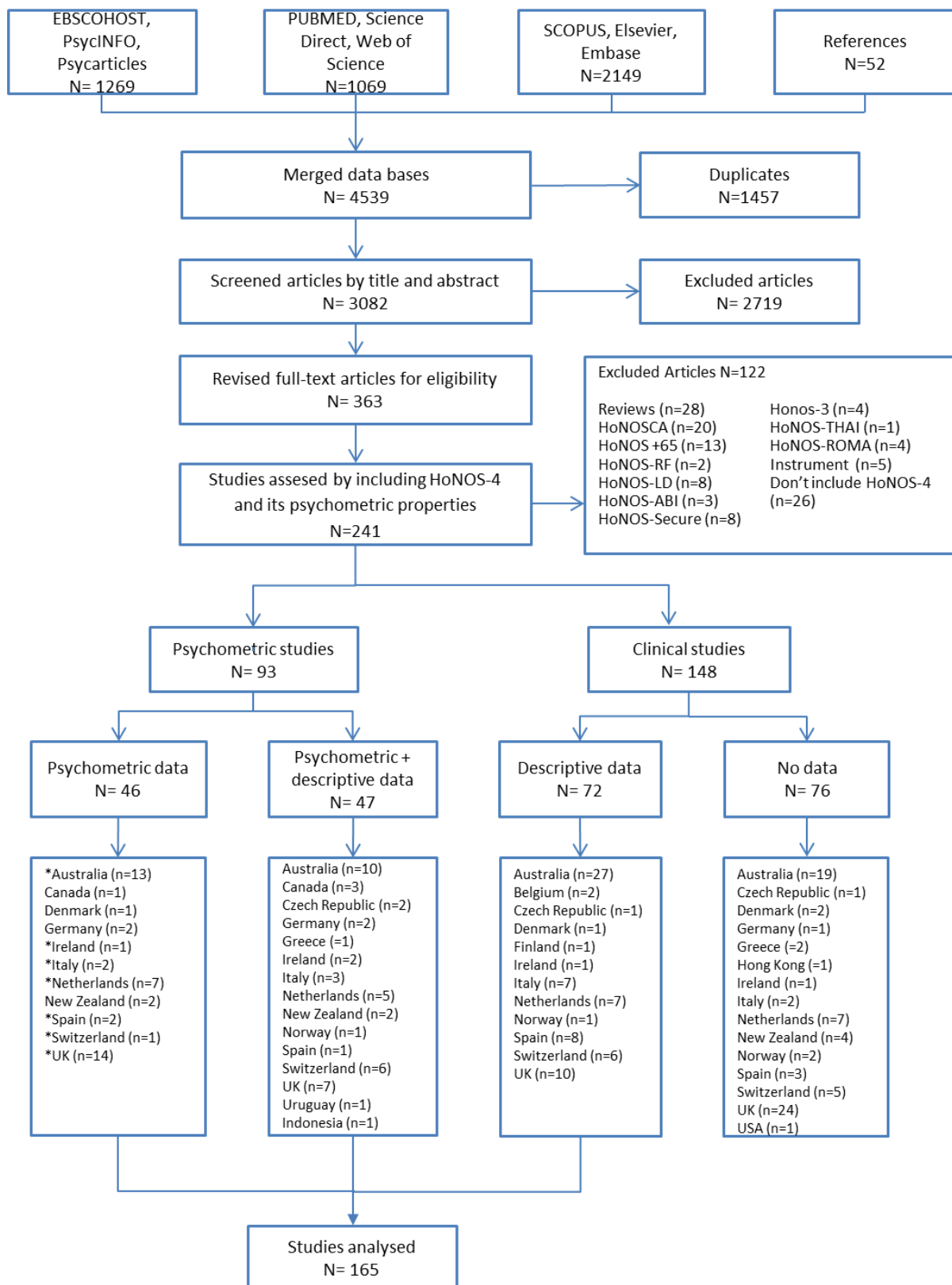
$$PC = \frac{(M_{Dis} \times DE_{Fun}) + (M_{Fun} \times DE_{Dis})}{DE_{Fun} + DE_{Dis}}$$

donde las M (medias) y DE (Desviaciones Estándar) pertenecen a valores de distribución normativos para muestras funcionales (Fun) y disfuncionales (Dis) existentes sobre la medida del fenómeno evaluado (Jacobson y Truax, 1991).

3.2.- Resultados

3.2.1.- Resultados de la estrategia de búsqueda

El proceso completo de identificación y selección de los estudios está prepresentado en el diagrama de flujo de la Figura 3.1. El proceso de revision sistemática ha detectado un total de 3082 artículos diferentes realizados entre 1998 y 2023 que potencialmente aludían en su



*HoNOS is used for the calculation of convergent validity with other instruments

Figura 3.1.- Diagrama de Flujo. Revisión Sistemática sobre el HoNOS

contenido a la escala HoNOS, y de los cuales han cumplido criterios de inclusión un total de 241 estudios realizados en 21 países provenientes de Europa (14), América (3), Asia (2) y Oceanía (2). De estos 241 estudios seleccionados, 93 tenían como objetivo el estudio de las características psicométricas del HoNOS, mientras que 148 utilizaban el instrumento para la valoración clínica de sus participantes. De este último grupo, se excluyeron 76 artículos que no aportaban datos concluyentes en cuanto a la caracterización clínica de la muestra (falta de datos sobre la desviación estándar, media de solo una de las subescalas, etc.), ni datos de caracterización psicométrica del HoNOS. Finalmente, se selecciona un total de 165 estudios independientes clasificados en tres categorías (Figura 3.1), que finalmente ofrecen evidencias que aglutinan datos psicométricos ofrecidos por 93 estudios (Anexo 1) y datos clínicos obtenidos de 119 estudios (Anexo 2). De estos estudios se extrajeron los datos para analizar y satisfacer los objetivos del estudio: datos psicométricos de validez y fiabilidad del HoNOS, y datos de la muestra clínica evaluada a través de la escala HoNOS (media, desviación típica y alpha de Cronbach).

De los 93 estudios que aportan datos psicométricos, el 47.3% provienen de Australia (n=23) y Reino Unido (n=21). De igual manera, en lo referente a los datos clínicos recogidos de 119 estudios, Australia es el país que más datos consistentes ofrece a través de la escala HoNOS (n=37), seguido de Reino Unido (n=17) y Países Bajos (n=12).

3.2.2.- Validez de contenido / aparente

En la revisión sistemática se han recogido tres estudios que evalúan la validez aparente, dos en Reino Unido en 1999 (Orrell, Yard, Handysides and Schapira) y en 2004 (Hunter, et al.) y en Australia en el 2009 (Burgess, et al.). Los tres reconocen la presencia de validez aparente a lo largo de los ítems de la escala. No se han encontrado más estudios posteriores que exploren o presenten datos de esta propiedad psicométrica.

3.2.3.- Validez estructural o factorial

En cuanto a la validez estructural del HoNOS, 17 estudios analizan este tipo de validez. Se recogen propuestas factoriales de cuatro factores (11 estudios), cinco factores (2 estudios), y un único factor (4 estudios). Por otra parte, un estudio propone la escala HoNOS como un instrumento unidimensional. Los datos recogidos de la validez factorial se reflejan en la Tabla 3.1.

Tabla 3.1.- Estructuras factoriales sugeridas para el HoNOS

Wing (1998)	Ítems	Estudios convergentes versión original* Preston (2000) Andreas (2010)	Trauer (1999) Eagar (2005)	McClelland (2000)	Newnhan (2009)	Andreas (2010)	Lovaglio (2011)	Lovaglio (2012)	Speak (2014)	Dogmanas (2014) William & Speak (2014)	Speak (2015)	Speak & Muncer (2016) Golay (2016)	Muncer (2016)	Smith (2017)
F1 – Problemas conductuales	It01 – Agresividad	1	1	1	1	1	1			1		4		1
	It02 – Autoagresiones	1	3	2	1	1	2			1		2	2	2
	It03 – Uso de sustancias	1	1	1	1	1	4			1		3		1
2 – Deterioro	It04 – Disfunción cognitiva	2	2	1	2	2	1	1		1		1		3
	It05 – Discapacidad Física	2	2	3	2	2	3			1		1		3
F3 – Problemas Clínicos	It06 – Alucinaciones/delirantes	3	5	2	2	3	1	1		1		4		1
	It07 – Depresión	3	3	1	3	3	2			1		2	2	2
	It08 – Otros Síntomas	3	3	4	3	3	2			1		2	2	1,2
F4 – Problemas sociales	It09 – Relaciones Sociales	4	3-4	1	1	4	1	1	1	1	1	3		4
	It10 – Funcionamiento general	4	4	1	2	4	1	1	1	1	1	1	1	3,4
	It11 – Problemas de residencia	4	4	1	4	4	1	1	1	1	1	3	1	4
	It12 – Problemas ocupacionales	4	4	1	4	4	1	1	1	1	1	1-3	1	4

Nota. Los factores de Preston, McClelland y Andreas son los mismos que los propuestos por Wing en el estudio original. Trauer define los factores de la siguiente forma: F1=Conducta, F2=deterioro, F3=Depresión, F4=Problemas sociales y F5=Alucinaciones. Newnhan los define: F1=Problemas conductuales o antisociales, F2=Deterioro, F3=Síntomas y F4=Socioeconómico. Speak (2016) los define: F1= Bienestar personal, F2= Bienestar emocional, F3=Bienestar social y F4= Alteraciones severas. Muncer los define: F1=Social, F2=Depresión. Smith los nombra: F1=diagnóstico de adicciones, conducta y salud mental, F2=Estabilidad emocional, F3=enfermedad cognitiva o física y cuidado personal y F4=Aspectos sociales. Lovaglio (2011): F1=Gravedad de la enfermedad, F2=Salud mental, F3=Problemas físicos, F4=Abuso de sustancias.

3.2.4.- Validez de criterio / convergente

Treinta y siete estudios aportan datos de validez convergente de HoNOS respecto a un total de 33 instrumentos de medida de constructos afines. Los instrumentos más usuales son la escala GAF (Global Assessment of Functioning) (n=9), BPRS (n=6), CANSAS y SCAN (n=4), y SCL-90-R y SBS (n=3). En la Tabla 3.2 se presentan los datos de integración de los índices de validez convergente con los instrumentos más utilizados.

Tabla 3.2.- Integración validez convergente instrumentos más usados

Instrumentos	Estudios	Promedio		Intervalo de Confianza		I ²	TAU	Q
		N	r _{pooled}	Inferior	Superior			
GAF	9	1818	-.62	-.47	-.77	98.30%	0.219	260.403
BPRS	6	3783	.63	.49	.78	97.96%	0.178	118.767
CANSAS	4	1292	.75	.68	.82	87.71%	0.067	25.000
SCAN	4	186	.57	.32	.83	84.46%	0.204	16.393
SBS	3	215	.44	.24	.64	70.25%	0.149	6.750
SCL-90-R	3	3435	.52	.35	.68	90.64%	0.135	10.543

Nota.- GAF: Global Assessment of Functioning Scale; BPRS: Brief Psychiatric Rating Scale; CANSAS: Camberwell Assessment of Need short appraisal schedule; SCAN: Schedules for Clinical Assessment in Neuropsychiatry; BSS: Beck Scale for Suicide Ideation; SCL90R: Symptom Checklist-90-R.

Todos los datos integrados son significativos ($p < .001$) con asociación positiva entre el HoNOS y las escala BPRS ($r = .62$), CANSAS ($r = .62$), SCAN ($r = .62$), SCL-90-R ($r = .62$) y SBS ($r = .62$), salvo con el instrumento GAF cuya asociación es negativa ($r = -.62$). El conjunto de evidencias ofrece una alta heterogeneidad (I^2), desde Los valores de integración presentan valores significativos que aportan convergencia entre HoNOS y los otros instrumentos; sin embargo, las evidencias presentan una alta heterogeneidad en sus efectos (ver Figura 2.2 para sesgo de publicación), que oscila entre un 70,25% (SBS) hasta un 98,3% (GAF).

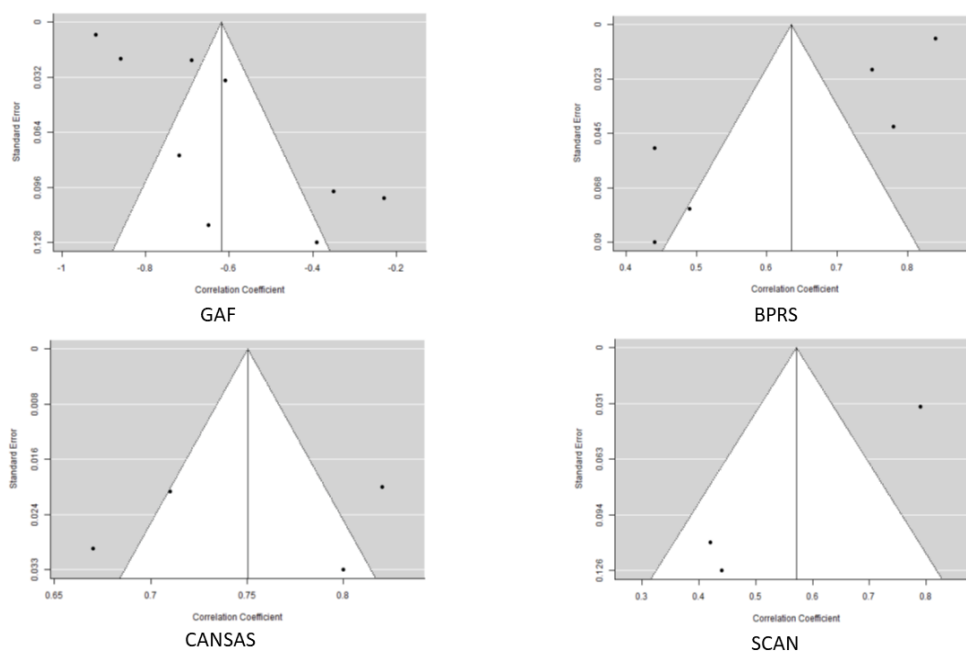


Figura 3.2.- Validez convergente GAF, BPRS, CANSAS y SCAN

3.2.5.- Validez discriminante

En lo referente a la validez discriminante de la escala, encontramos tres estudios que aportan datos de la escala HoNOS. Dos de ellos, (Dogmanas, 2014) y (Gallagher & Teesson, 2000), muestran una correcta actuación del HoNOS discriminando entre pacientes con diferentes niveles de severidad y discapacidad (disability), ya que recogen diferencias estadísticamente significativas entre las puntuaciones obtenidas por un grupo de pacientes internos y pacientes de servicio ambulatorio. El tercer estudio, (McClelland, et al., 2000), recoge los resultados de los grupos de pacientes, los cuales obtuvieron la mayor puntuación en el ítem HoNOS más relevante para su trastorno, lo que indica que la escala tiene cierta capacidad para discriminar entre los grupos diagnósticos.

3.2.6.- Validez predictiva

En lo referente a la validez predictiva, ésta es analizada por siete estudios. Cinco ofrecen datos positivos sobre este tipo de validez (McClelland, et al., 2000; Jury, et al., 2018; Golay, et al., 2016; Page, Hooke and Rutherford, 2001; Kisely et al, 2010). Se encuentran resultados de validez predictiva del HoNOS en variables como el tiempo de hospitalización, la probabilidad de readmisión psiquiátrica, el tiempo para la readmisión, contacto con el servicio, duración de la estancia, resultados previstos y permanencia en el servicio (Golay, et

al., 2016; Page, Hooke and Rutherford, 2001; Kisely et al., 2010); por otra parte, el estudio (Jury, et al., 2018) arroja datos de validez predictiva sobre el riesgo de reclusión/aislamiento. Por último, el estudio (McClelland, et al., 2000) concluyó que las puntuaciones en determinados ítems de la escala podrían predecir el grupo de diagnóstico de los pacientes, en concreto en pacientes con síntomas de psicosis, depresión y problemas con la alimentación, el sueño y el estrés.

Los estudios que no encontraron datos que apoyaran la validez predictiva fueron (Goldney, Fisher and Walmsley, 1998; Preti et al., 2012). El primero de ellos, realizado por Goldney, et al. (1998) no arrojó resultados significativos en cuanto a la variable ‘tiempo de hospitalización’. El segundo, realizado por Preti et al. (2012), tampoco encontró validez predictiva entre la puntuación obtenida y la remisión tanto en pacientes con primer episodio psicótico (first episode psychosis - FEP) como en pacientes psicóticos de alto riesgo (ultra high risk - UHR).

3.2.7.- Validez intercultural

La escala HoNOS ha sido adaptada en 17 países distintos al Reino Unido (país de la versión original) y traducida a 10 idiomas. En el proceso de adaptación de la versión italiana (Lora et al., 2001) y la versión tailandesa (Phuaphanprasert, et al., 2007), el número de ítems se ha visto modificado respecto a su versión original de 12 ítems (Wing et al., 1998). La versión en Italia de la escala HoNOS aumentó a 18 ítems y en Tailandia se redujo a diez. En ambas versiones también se constata la idoneidad psicométrica del instrumento.

3.2.8.- Estudios de adaptación y/o traducción a nivel internacional

Son cuatro los países que usan la versión original en inglés de UK y que muestran su adecuación psicométrica a la población objeto de estudio. Estos son: Australia (Trauer, 1999), Nueva Zelanda (Eagar, et al., 2005), Noruega (Øiesvold, Bakkejord and Sexton, 2011), Irlanda (Browne, Doran and McGauran, 2000). Países como EEUU (Tucker, et al., 2018), Bélgica (Lorant, et al., 2019; Wyngaerden, et al., 2019), Finlandia Leijala, et al., 2021), Indonesia (Anjara, et al., 2019) y Hong-Kong (Hariman, et al., 2020) hacen uso de la versión inglesa, pero no hay evidencia de estudios de adecuación psicométrica en esos países.

La escala HoNOS fue traducida al español en 1999 por el equipo de Uriarte et al. (1999); sin embargo, no se han encontrado estudios de adecuación psicométrica en muestra española.

Existe una adaptación de esta versión en población uruguaya (Dogmana, 2014), con una leve adaptación lingüística en dos ítems del glosario de la escala.

Por último, se recogen estudios de Italia (Lora et al., 2001), Países Bajos (Mulder, et al., 2004), Dinamarca (Bech, et al., 2003), Alemania (Andreas, et al., 2007), Canadá (Lauzon, et al., 2001), República Checa (Pěč, et al., 2009), Tailandia (Phuaphanprasert, et al., 2007) y Grecia (Mavreas and Fotiadou, 2010) que traducen la escala a su lengua materna (italiano, holandés, danés, alemán, francés, checo, tailandés y griego, respectivamente) y la adaptan psicométricamente en la población. Por su parte, Suiza (Theodoridou, 2011) usa la versión en alemán HoNOS-D para adaptarla en población suiza.

3.2.9.- Sensibilidad al cambio

En lo referente a la capacidad de la escala para detectar cambios en las puntuaciones, 19 estudios recogen resultados significativos acerca de la sensibilidad del HoNOS en detectar cambios en el tiempo, obteniendo tamaños del efecto moderados-altos ($d = .69$) pre-post alta (Audin, et al., 2001). El estudio de Newnham, Harwood, y Page (2009) señala que aumenta la sensibilidad al cambio con el modelo de cuatro factores modificado en tres de las cuatro subescalas.

Por otra parte, uno de los estudios que evalúa la sensibilidad al cambio en diferentes entornos (Trauer, et al., 1999), concluye que la escala es eficaz para detectar cambios a corto plazo en servicios agudos, pero no obtiene resultados significativos en entornos comunitarios.

3.2.10.- Fiabilidad

En lo referente a la consistencia interna, se ha realizado la integración metaanalítica de los índices Alpha de Cronbach aportados por los estudios recogidos ($k = 33$) que aportan el Alpha de Cronbach, en un rango que oscila entre .45 a .86. (Tabla 3.3).

El valor promedio de consistencia interna del total de estudios considerados ha sido de .77 (IC95% .60 a .78), que puede considerarse como una fiabilidad moderada alta, no obstante, se aprecia una alta heterogeneidad entre estudios ($I^2 = 98.30\%$; $TAU = .09$; $Q = 4010.87$), que también resulta apreciable en el diagrama de representación de efectos (Forest Plot – Figura 3.3), donde los estudios de Newnham et al (2009), Smith et al (2017), que presentan resultados muy consistentes (IC muy estrechos) junto a los de Machova et al (2023) y Tennneij et al (2009), con resultados menos consistentes (IC muy anchos), pero alejados todos ellos del promedio del efecto encontrado en el conjunto. Hacer notar que estos cuatro

Tabla 3.3.- Integración meta analítica de los datos de fiabilidad recogidos

País	Tamaños		Valores de Alpha (α)				Heterogeneidad		
	K Estudios	N muestra	Min	Max	Efecto Promedio	IC 95%	I ² (%)	TAU	Q
Reino Unido	5	95980	.52	.86	.80	.57 a .80	99.67	.13	3358.65
Nueva Zelanda	1	12576	-	-	.80	-	-	-	-
Australia	7	9961	.45	.76	.53	.57 a .74	97.64	.11	584.61
España	1	281	-	-	.70	-	-	-	-
Uruguay	1	402	-	-	.69	-	-	-	-
Italia	4	12233	.70	.78	.72	.71 a .74	51.20	.01	6.81
Países Bajos	8	1891	.48	.78	.72	.66 a .76	84.97	.06	29.49
Finlandia	0	-	-	-	-	-	-	-	-
Dinamarca	0	-	-	-	-	-	-	-	-
Noruega	1	118	-	-	.67	-	-	-	-
Alemania	2	3586	.62	.73	.63	-	-	-	-
Suiza	1	53	-	-	.50	-	-	-	-
Bélgica	0	-	-	-	-	-	-	-	-
Canadá	1	194	-	-	.70	-	-	-	-
Irlanda	0	-	-	-	-	-	-	-	-
Chequia	2	97	.52	.79	.62	-	-	-	-
Indonesia	0	-	-	-	-	-	-	-	-
TOTAL	34	137372	.45	.86	.77	.60 a .78	98.30	.09	4010.874

Notas.-

estudios que se apartan en mayor medida del perfil global, pertenecen a cuatro países diferentes, por lo que habría que estudiar la causa de su aportación a la heterogeneidad en las características de la población estudiada u otros factores de influencia.

Hacer notar, por otro lado, como los países que mayores valores promedio de alpha alcanzan (.80) son Reino Unido y Nueva Zelanda, siendo asimismo los que mayor número de muestran presentan.

Respecto a otros criterios de fiabilidad analizados se ha explorado la Fiabilidad intrasujeto (Coeficiente de correlación Intraclass (ICC) y test-retest) en 15 estudios donde se aprecia un rango de valores entre .29 y .92; y la Fiabilidad inter-jueces (Coeficiente Kappa), donde se han encontrado aportaciones de siete 7 estudios aportan datos con un rango entre .42 y .90.

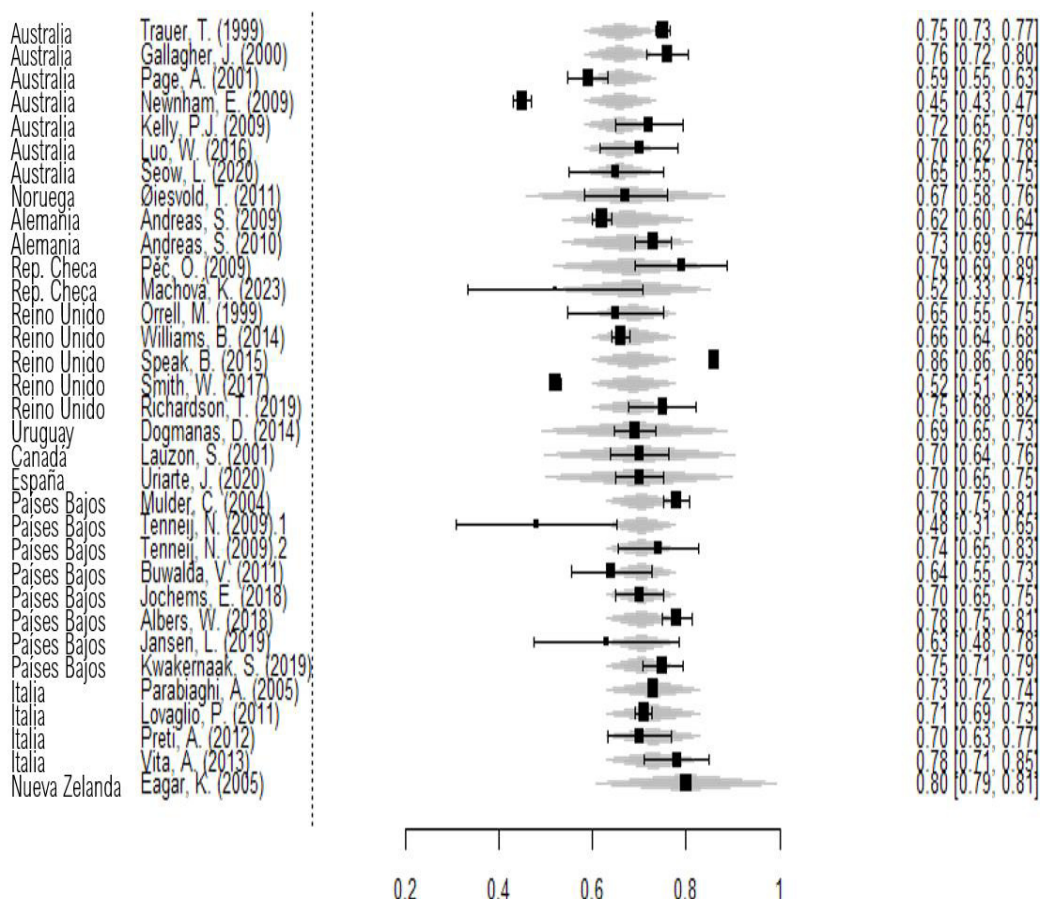


Figura 3.3.- Integración meta-analítica de valores alpha del HoNOS

3.2.11.- Interpretabilidad

No se han encontrado estudios que aporten datos de interpretabilidad del HoNOS.

3.2.12.- Puntos de corte y Error estándar de medida

Cuatro estudios proporcionan puntos de corte para la interpretación de la escala HoNOS. El primero es el estudio de validación y adaptación de la escala HoNOS en Uruguay (Dogmanas, 2014). Aplicando la metodología de cambio fiable de Jacobson y Truax (1991) obtienen un punto de corte de 7.3 para la diferenciación entre expresión sintomática normativa y el grado de severidad y un valor de Puntuación Mínima de Cambio (PMC) o Índice de Mejoría Fiable (IMF) de 6.07. El segundo estudio, con muestra clínica de Australia, (Maxwell, et al, 2018) ofrece un valor para el punto de corte de 6.38 y un valor de PMC de 4.59. El tercero es un estudio clínico con muestra de Países Bajos (Jansen, et al., 2019), ofrece un punto de corte de 14. El último estudio (Parker, et al., 2020) ofrece datos de referencia,

Tabla 3.4.- Integración de datos clínicos por países y para el total de estudios revisados

País	Tamaños		Media		Variabilidad y Error					Puntos Corte		Punt. Cambio	
	K	N	min	max	M _{Pro}	IC 95% M _{Pro}	DE _{Pro}	SE	SEM	PC _{JT}	PC _{1sd}	SE-Dif	PMC
Reino Unido	17	71859	7.05	15.17	11.43	11.38 a 11.48	6.87	0.03	3.03	0.10	4.56	4.29	7.08
Nueva Zelanda	2	70	8.00	13.87	11.77	10.24 a 13.3	6.57	0.78	2.93	0.93	5.20	4.15	6.86
Australia	37	112204	6.46	19.30	14.03	13.99 a 14.07	6.65	0.02	4.57	3.10	7.38	6.46	10.66
España	9	3289	9.05	20.13	14.58	14.38 a 14.78	5.89	0.10	5.89	4.87	8.69	8.32	13.73
Uruguay	1	402	-	-	7.60	7.19 a 8.01	4.20	0.21	2.34	0.67	3.40	3.31	5.46
Italia	10	14542	8.70	21.90	9.62	9.52 a 9.72	6.11	0.05	3.19	-0.46	3.51	4.52	7.45
Países Bajos	12	23043	7.15	20.60	12.31	12.23 a 12.39	6.48	0.04	3.35	1.63	5.84	4.74	7.83
Dinamarca	1	357	-	-	11.01	10.44 a 11.58	5.48	0.29	5.48	1.97	5.53	7.75	12.79
Finlandia	1	276	-	-	7.20	6.67 a 7.73	4.50	0.27	4.50	-0.22	2.70	6.36	10.50
Noruega	2	1035	9.49	12.06	11.77	11.04 a 12.49	11.9	0.37	6.84	-7.87	-0.14	9.67	15.96
Alemania	2	5087	16.56	16.99	16.83	16.69 a 16.97	5.02	0.07	3.05	8.54	11.80	4.31	7.11
Suiza	12	58482	1.14	24.60	19.15	19.09 a 19.21	7.66	0.03	7.66	6.51	11.49	10.84	17.88
Bélgica	2	1787	11.6	12.00	11.91	11.6 a 12.22	6.78	0.16	6.78	0.73	5.13	9.59	15.82
Grecia	1	544	-	-	16.28	15.78 a 16.78	5.96	0.26	5.96	6.45	10.32	8.43	13.91
Canadá	3	5011	10.00	12.20	10.12	9.96 a 10.28	5.52	0.08	5.52	1.01	4.60	7.80	12.90
Irlanda	3	650	4.43	11.80	6.31	5.88 a 6.74	5.64	0.22	5.64	-2.99	0.67	7.98	13.16
Chequia	3	357	2.05	11.31	4.18	3.68 a 4.67	4.77	0.25	2.94	-3.70	-0.60	4.16	6.86
Indonesia	1	287	-	-	5.45	4.86 a 6.04	5.13	0.30	5.13	-3.01	0.32	7.25	11.98
TOTAL	119	299299	1.14	24.6	13.98	13.11 a 3.17	7.46	0.014	3.60	1.67	6.52	5.08	8.38

Notas.- K: nº de estudios; N: tamaño de muestra acumulada; Min-Max: Medias mínima y máxima en los estudios de un determinado país; Mpro: Media promedio integrada del país; IC95%: Intervalo de confianza del 95% de la Mpro; DEpro: Desviación integrada por país; SE: Error de Estándar; SEM: Error estándar de medida; PCJT: Puntuación de Corte por el Método de Jacobson y Truax; PC1SD: Punto de Corte basado en una desviación estándar por debajo de la media; SEDif: Error estándar de la diferencia; PMC: Puntuación Mínima de Cambio.

también de Australia, con un valor de 13 como punto de corte y una Puntuación Mínima de Cambio de 6.14. Estos dos últimos estudios ofrecen valores apreciablemente mayores que los propuesto por los dos primeros estudios.

En el presente estudio proponemos puntos de corte a través de la integración meta analítica de los datos proporcionados en los estudios analizados (muestra, media, desviación estándar) para cada país. Además de la PMC (Puntuación Mínima de Cambio), ofrecemos dos medidas como referencia de punto de corte (PC), una siguiendo la formulación de Jacobson y Truax y la otra restando una desviación estándar a la media integrada. Estos datos quedan recogidos en la Tabla 3.4.

3.3.- *Discusión*

Se analizaron un total de 165 estudios, 93 de ellos aportaban datos de caracterización psicométrica de la escala HoNOS, y 119 aportaban gravedad clínica. Reino Unido y Australia representan el 49,1% de la muestra de estudios analizados (n=81), debido a que ambos países integran la escala HoNOS en protocolos de evaluación en la práctica clínica diaria en los servicios de salud nacionales (Speak and Muncer, 2015; Smith, et al., 2017; Penington, et al., 2022; Burgess, Pirkis, and Coombs, 2009; Coombs, 2014; Segal, Rimes, and Hayes, 2019). Si bien Nueva Zelanda (Eagar, Trauer, and Mellsop, 2005; n=12576; y el estudio de Lai, et al., 2018; n=10727) y Países Bajos (De Beurs, et al., 2018; n=16671) también incluyen la escala HoNOS en su protocolo de valoración clínica, los estudios recogidos no ofrecen datos de caracterización psicométrica concluyentes. Por otra parte, países como Italia y los Países Bajos (Italia art 7 n=9817) muestran en el meta-análisis un gran conjunto de datos clínicos (N total Italia=13923; N total Países Bajos=21291) lo que nos hace suponer que la escala HoNOS es administrada de forma sistemática y se podría encontrar dentro de los protocolos de los diferentes centros sanitarios. En conjunto, los datos obtenidos en la revisión sistemática como integración del conjunto de estudios internacionales realizados sobre la escala HoNOS muestran un comportamiento adecuado para su utilización en la valoración y monitorización de procesos de enfermedad mental grave.

En lo referente a la validez de contenido o aparente, en la revisión sistemática se han recogido tan solo dos estudios que evalúan este tipo de validez, en 1999 (Orrell, et al) y en 2004 (Hunter, et al), ambos en Reino Unido. No se han encontrado más estudios posteriores que exploren o presenten datos de esta propiedad psicométrica, por lo que puede deducirse que la validez aparente es una propiedad que se da por hecho en la escala HoNOS.

En el caso de España, no se ha localizado ningún estudio que examine el grado de adecuación psicométrica de la versión adaptada idiomáticamente. Sí se han encontrado estudios que hacen un uso clínico de la escala HoNOS (Gisbert, and Cid Colom, 2010; Frades, et al., 2011; Osta Samanes, et al., 2012; Escalada-Hernández, et al., 2014; Garay Arostegui, et al., 2014; ; Escalada-Hernández, et al., 2015; Escalada-Hernández, and Marín-Fernández, 2015; Ribas, et al., 2016; Uriarte, et al., 2024) (n=9) así como otros estudios que aportan datos de validez convergente de la escala HoNOS (Lascorz, et al., 2016; Vargas, et al., 2013), a pesar de no estar validada en población española (n=2). Esto demuestra el gran uso del HoNOS en España tanto con propósitos psicométricos como clínicos. Sin embargo, es

importante que un instrumento de evaluación muestre buena fiabilidad, validez y sensibilidad al cambio. Marshall et al. (2000) demostró empíricamente que los resultados de los estudios difieren cuando estos emplean instrumentos validados respecto a aquellos en los que usan instrumentos sin validar (Mokkink et al, 2010a). Teniendo esto en cuenta, la escala HoNOS, a pesar de no estar validada, es un instrumento de referencia en la práctica clínica diaria en diferentes servicios sanitarios de España (Uriarte, et al., 1999) por lo que consideramos necesaria su validación tanto para examinar su convergencia con los resultados internacionales como para obtener resultados fiables en los estudios de carácter clínico.

OBJETIVOS

Capítulo 4

Objetivos

4.1.- Objetivo(s)

Dada la falta de estudios de validación de la escala HoNOS para su uso en población española encontrados en la revisión presentada en el capítulo anterior, el objetivo principal de la presente tesis es obtener datos del comportamiento psicométrico de la escala HoNOS en población española. La presentación de evidencias de validez y fiabilidad de la escala contribuirá a avalar la pertinencia y oportunidad de su uso en contextos clínicos para el apoyo en el diagnóstico (valoración de la gravedad sintomática de forma global o por sub-escalas) y, sobre todo, la posibilidad de utilizar el HoNOS como un instrumento de valoración/evaluación de resultados en la práctica clínica rutinaria.

La satisfacción de este objetivo general se pretende alcanzar por medio de la realización de cuatro estudios cada uno de los cuales presentará un objetivo específico. Se enuncian a continuación en función del orden del estudio realizado:

Objetivo Estudio 1.- Ofrecer evidencias de adecuación psicométrica de la escala HoNOS en un estudio piloto en una muestra aleatoria (n= 281) de personas con Enfermedad Mental Grave (PEMG) atendidas en el servicio de la Red de Salud Mental de Bizkaia (RSMB).

Objetivo Estudio 2.- Ofrecer nuevas evidencias de adecuación psicométrica de la escala HoNOS en población de Personas con Enfermedad Mental Grave (PEMG) atendidas en el servicio de la Red de Salud Mental de Bizkaia (RSMB) (n= 3108). En este caso, se trata de reunir las evidencias de

fiabilidad (consistencia interna y fiabilidad temporal-test retest) y validez (constructo, convergente, discriminante y predictiva) de la escala HoNOS obtenidas tanto a través de la Teoría Clásica de los Test (TCT) como de la Teoría de Respuesta al Ítem (TRI).

Objetivo Estudio 3.- Ofrecer baremos para la clasificación y evaluación de resultados en PEMG de la RSMB y puntos de corte y valores de referencia de cambio clínico a través de la Puntuación Fiable de Cambio atendiendo a las variables de segmentación sexo, servicio de atención clínica (de forma independiente o combinada).

El cumplimiento de los dos primeros objetivos permitirá decidir si la escala HoNOS constituye una herramienta válida y fiable para su utilización en la valoración/evaluación de sintomatología en casos con Enfermedad Mental Grave, y su posible aplicación como instrumento de medida de resultados en Salud Mental.

El tercer objetivo permitirá proporcionar a clínicos y servicios asistenciales los procedimientos útiles para interpretar los resultados individuales obtenidos por el usuario, tanto en la puntuación total como en las dimensiones y con ello facilitar la toma de decisiones clínicas.

EVIDENCIAS / ESTUDIOS

Capítulo 5

Estudio 1: Evidencias de Adecuación Psicométrica de la Versión Española de la Escala HoNOS³

5.1.- Objetivo

Ofrecer evidencias de adecuación psicométrica de la escala HoNOS en un estudio piloto en una muestra aleatoria (n= 281) de personas con Enfermedad Mental Grave (PEMG) atendidas en el servicio de la Red de Salud Mental de Bizkaia (RSMB).

5.2.- Método

5.2.1.- Contexto de realización de los estudios

La Red de Salud Mental de Bizkaia (RSMB) nace en el año 2010. Bajo una única dirección integra todos los recursos ofertados desde la atención en Salud Mental Extrahospitalaria (Centros de Salud Mental -CSM-, Hospitales de Día -HD-, Tratamiento Asertivo Comunitario -TAC-) y los tres hospitales psiquiátricos de Bizkaia: Hospital de Bermeo, Hospital de Zaldibar y Hospital de Zamudio.

³ Artículo publicado que puede encontrarse en el anexo 9. Referencia: Uriarte, J. J., Iglesias, N., Penas, P., Moreno-Calvete, M. C., Álvarez-González, A. y Iraurgi, I. (2024). Evidencias de Adecuación Psicométrica de la Versión Española de la Escala HoNOS. *Revista Iberoamericana de Diagnóstico y Evaluación – e Avaliação Psicológica. RIDEP*, 72(2), 145-156.

Desde su creación, desarrolla un modelo de atención comunitario basado en tres objetivos: el crecimiento de la capacidad de atención en el ámbito comunitario y cercano a las personas que precisan atención; la reducción a lo indispensable de las estancias hospitalarias; e intensificar la coordinación con los servicios sanitarios, sociales y sociosanitarios para prestar una atención transversal e integrada a la persona con enfermedad mental.

El núcleo de la atención es el Centro de Salud Mental. Junto a él, los Hospitales de Día y los Equipos de Tratamiento Asertivo Comunitario prestan atención a los pacientes en su ambiente habitual y ayudan a las personas con enfermedad mental grave a mantenerse en el entorno de la comunidad. Los servicios disponibles se completan con las unidades de hospitalización de Bermeo, Zaldibar y Zamudio.

De forma rutinaria, cuando una persona contacta con la RSMB por primera vez y también en posteriores seguimientos (generalmente cada 4 o 6 meses) se aplica el protocolo de valoración, el cual está formado por la escala HoNOS, el Índice Clínico Global (ICG) y la escala EEAG. Estas evaluaciones se incluyen en el historial médico Osabide Global.

5.2.2.- Muestra

En el contexto de un estudio de evaluación de la recuperación de personas con enfermedad mental grave en tratamiento en la Red de Salud Mental de Bizkaia (RSMB), se obtuvo una muestra representativa de 281 participantes seleccionados de un total de 1949 usuarios censados en la RSMB asumiéndose un error de estimación de 5.1% para un nivel de confianza del 95%. Para la selección se llevó a cabo una estratificación en función del sexo, edad y tipo de centro asistencial con asignación aleatoria a cada estrato. Como criterios de inclusión se consideraron: ser mayor de 18 años y la pertenencia a la RSMB. Como criterios de exclusión: la no obtención del consentimiento informado, dificultades en la comunicación y presentar un estado clínico que no permita la colaboración del o la participante.

Los participantes presentaban una edad media de 48.8 años (DE=11.2) con edades comprendidas entre los 23 y 81 años, siendo preferentemente hombres (n=174; 61.9%). Respecto a su distribución entre los distintos programas de tratamiento, 164 recibían asistencia en los Centros de Salud Mental (CSM), 74 en Hospitales de Día (HD), 22 en el Tratamiento Asertivo Comunitario (TAC) y 21 en Hospitales Psiquiátricos (HP). El diagnóstico más frecuente fue de esquizofrenia (54.2%), seguido del trastorno bipolar (12.3%) y el trastorno esquizoafectivo (9.9%), con una media de 17.85 años en tratamiento

(DE=8.64, rango=2-45). Un 76.9% de los participantes han sufrido alguna hospitalización a lo largo de su evolución con una media de 5.16 hospitalizaciones (DE=7.89). Un 78.5% de los casos tenían acreditado un grado de minusvalía (media=65.06, DE=9.82), asimismo un 77% no tenían reconocido ningún grado de dependencia. En relación a la convivencia, un 66.1% viven con sus familias, un 1.4% autónomamente y un 7.5% en alojamientos supervisados.

5.2.3.- Instrumentos

Health of the Nation Outcome Scale – HoNOS (Wing et al., 1998). Se trata de una escala de 12 ítems de aplicación por parte del personal clínico que a partir de una escala de intensidad de 0 a 4 puntos (0=sin problema, 1=problema menor que no requiere intervención, 2=problema leve pero claramente presente, 3=problema de moderada gravedad y 4=problema grave o muy grave) se valora la gravedad del problema que tiene el paciente respecto a cada uno de los ítems. Permite obtener una puntuación global a partir de la contribución de todos los ítems, y cuatro medidas específicas correspondientes a las dimensiones teóricas propuestas por los autores. Existe también la posibilidad de utilizar cada ítem como un indicador único de la manifestación a la que hace referencia (Bebbington et al., 1999), si bien esta opción es poco utilizada y recomendada (Speak & Muncer, 2015). Un reciente estudio de revisión sistemática de las evidencias de validez del HoNOS (Iglesias et al., 2019) ofrece datos de su consistencia y valores de precisión de medida (fiabilidad) con coeficientes que oscilan entre .37 y .86, siendo la fiabilidad promedio integrada por meta-análisis de .81. En el presente estudio se ha utilizado la adaptación al castellano utilizada por el grupo de Uriarte y colaboradores (1999).

Escala de Evaluación de la Actividad Global – EEAG (Endicott et al., 1976; Bobes et al., 2002). Es una escala de valoración de 0 a 100 puntos diseñada para que un clínico evalúe el nivel general de actividad de un paciente, en la cual una mayor puntuación significaría un mejor funcionamiento. Este instrumento constituye la información del eje V del diagnóstico multiaxial del DSM (APA, 2004).

Escala de Impresión Clínica Global – ICG (Guy, 1976; 2004), en concreto la sub-escala que evalúa la gravedad del cuadro clínico (CGI-SI), se trata de una escala Likert donde el clínico asigna un valor de 0 (no evaluado) a 7 (extremadamente enfermo) donde se califica la gravedad global del padecimiento del paciente.

Escala analógico visual del cuestionario de Salud Euro-QoL5D – EQ5D (EuroQol Group, 1990; Herdman et al., 2001). Se trata de una escala analógico-visual en forma de termómetro donde el participante marca “su estado de salud hoy”, siendo 0 el peor estado de salud que pueda imaginarse y 100 el mejor.

5.2.4.- Procedimiento

De manera protocolizada, el programa de Trastorno Mental Grave de la RSMB tiene implementado un procedimiento de valoración del estado de gravedad y evolución en el programa de los usuarios que acceden al mismo. La valoración de cada sujeto en el momento de evaluación que le corresponde es realizada de forma colegiada por el equipo que le asiste a través de los instrumentos HoNOS, ICG y EEAG, habiéndose solicitado al participante la cumplimentación del Euro-QoL5D. Para el presente estudio se tomó como base de análisis la última valoración realizada de cada participante seleccionado para el estudio. En todo momento el procedimiento de evaluación sigue los estándares de comportamiento ético en la investigación y cuenta con la aprobación del CEIC (Comité Ético de Investigación Clínica de Euskadi - Osakidetza).

5.2.5.- Aspectos éticos

El proyecto de investigación obtuvo la aprobación del Comité Ético de Investigación Clínica de Euskadi (CEIC-Euskadi) para su desarrollo, recogido en el acta 8/2012 fechada 10 de octubre de 2012.

Todos los usuarios que fueron seleccionados como posibles participantes del estudio y cumplían los criterios de inclusión, recibieron información acerca del objetivo del estudio y de cuál sería su implicación en el mismo una vez hubiesen firmado el consentimiento informado. Todos los/as posibles candidatos/as, aceptasen o no finalmente participar, fueron contactados/as por su responsable sanitario de la RSMB quien le transmitía información básica sobre el proyecto y solicitaba su participación. Para complementar la información recibida por el clínico correspondiente se les entregaba una hoja de información sobre el proyecto (anexo 5) que debían leer, preguntar cualquier duda que les pudiese surgir, y posteriormente, si estaban de acuerdo y querían participar en el estudio, pasaban a rellenar y firmar el consentimiento informado (anexo 6). Dada la gravedad del colectivo, aquellos usuarios que tuviesen un tutor legal (incapacidad legal o curatela sanitaria), se informaba del proyecto de investigación tanto al usuario como al tutor legal (anexo 7), pero en este caso era

este último quien firmaba el consentimiento informado (anexo 8). En todos los casos el consentimiento era firmado por duplicado, una copia destinada al usuario junto a la hoja informativa, y una segunda copia quedaba en poder del equipo investigador encargado de custodiarla.

La información de las personas participantes en el estudio se trató de forma anónima según el procedimiento de la Ley Orgánica 15/1999 de 13 de diciembre de protección de datos de carácter personal. Este proyecto de investigación respetó los principios fundamentales establecidos en la Declaración de Helsinki, así como los requisitos establecidos por la legislación vigente en el ámbito de la investigación médica. Con el objetivo de preservar la integridad y confidencialidad de los datos del estudio, a los participantes se les asignó un código de identificación que permitía garantizar el anonimato de la información recogida durante el estudio, y evitase vinculación alguna con los datos de la historia clínica del participante.

5.2.6.- Análisis de datos

Para el análisis de los ítems de HoNOS se utilizaron la frecuencia (n), porcentaje, media (M), desviación estándar (DE), asimetría (As), correlación entre el ítem con el resto de la escala (r), Alfa de Cronbach (α) si se elimina elemento y Alfa de Cronbach para el total y las dimensiones. Para el análisis de asociación entre variables se utilizó el coeficiente de correlación de Pearson (r).

Un primer análisis factorial exploratorio se llevó a cabo con el programa Factor 10 (Lorenzo- Seva & Ferrando, 2019). Se verificó la idoneidad de la matriz de correlación a través de la prueba Kaiser-Meyer-Olkin y la prueba de esfericidad de Bartlett. El test de Análisis Paralelos (PA; Parallel Analysis) y el método de Hull se utilizaron como criterios de extracción para el número de factores. Asimismo, se analizó la normalidad multivariada con la prueba de Mardia. Los análisis factoriales confirmatorios (CFA) se llevaron a cabo con el programa EQS 6.1 (Bentler, 2006), utilizando el método de estimación de máxima verosimilitud. Se probaron todos aquellos modelos encontrados en la revisión teórica, incluyendo las estructuras bifactoriales de estos. Para evaluar el nivel de bondad de los distintos modelos se han utilizado los siguientes indicadores: la prueba chi- cuadrado χ^2 de igualdad entre matriz y grados de libertad (g.l.), el criterio de información de Akaike (AIC), bondad del índice de ajuste (GFI), el índice ajustado no normado (NNFI), el índice de ajuste comparativo (CFI), el residuo cuadrático medio estandarizado (SMSR), el error cuadrático

medio de aproximación (RMSRA) junto con su intervalo de confianza (LI-LS) del 90%. Se utilizaron los siguientes criterios para evaluar la bondad de ajuste: los valores de GFI, NNFI y CFI deben superar el valor de .90 (cuanto mayor sea el valor, mayor será el ajuste) y los valores de SMR y RMSEA deberán ser menores o iguales a .05, aunque también pueden asumirse valores de hasta .07.

Para el contraste de medias se utilizaron pruebas de análisis de varianza (F), que en caso de falta de homocedasticidad entre grupos se aplicó la corrección del método robusto basado en la prueba de Brown-Forsythe. En caso de significación estadística de la prueba F, el contraste de medias entre grupos se realizó mediante pruebas post-hoc basados en el método de Scheffe.

Tabla 5.1.- Estadísticos descriptivos y medidas de tendencia central de los ítems y dimensiones del HoNOS

	Distribución de respuestas					Estadísticos				Consistencia		
	0	1	2	3	4	M	DE	CV	As	r	α	h^2
1	74.0	13.5	9.3	2.5	.7	0.42	0.82	195.24	2.14	.32	.69	.22
2	96.1	1.1	1.1	.7	1.1	0.10	0.53	530.00	6.07	.32	.69	.21
3	76.5	10.0	8.9	3.9	.7	0.42	0.86	204.76	2.05	.20	.70	.15
4	29.9	35.2	26.7	8.2	.0	1.13	0.94	83.19	0.34	.27	.69	.16
5	39.1	21.0	25.6	11.4	2.8	1.18	1.15	97.46	0.55	.17	.71	.12
6	41.3	22.8	24.6	9.6	1.8	1.08	1.10	101.85	0.62	.22	.70	.13
7	43.1	27.8	21.0	7.1	1.1	0.95	1.01	106.32	0.76	.31	.69	.21
8	43.1	14.6	24.2	15.7	2.5	1.20	1.22	101.67	0.47	.38	.68	.27
9	13.9	20.6	38.4	23.8	3.2	1.82	1.05	57.69	-0.19	.56	.65	.43
10	28.1	23.1	30.6	16.7	1.4	1.40	1.11	79.29	0.15	.65	.63	.52
11	66.5	12.5	11.4	3.9	5.7	0.70	1.17	167.14	1.63	.34	.69	.27
12	54.8	16.4	17.4	6.8	4.6	0.90	1.19	132.22	1.12	.38	.68	.27
Dimensiones												
Problemas Conductuales						0.94	1.49	158.51	1.81	.38	.35	.43
Deterioro						2.31	1.52	65.80	0.20	.30	.10	.29
Problemas Clínicos						3.23	2.31	71.52	0.72	.37	.45	.42
Problemas Sociales						4.82	3.21	66.19	0.69	.54	.68	.67
Total						11.30	5.98	52.92	0.55		.70	

Nota. 0=Sin problema, 1=problema menor que no requiere intervención, 2=problema leve pero claramente presente, 3=problema de moderada gravedad o 4=problema grave o muy grave; M=media, DE=Desviación estándar, CV=Coeficiente de Variación, As=Asimetría, α =Alfa de Cronbach si se elimina el elemento y para los totales de la escala se informa de alfa de Cronbach obtenida, h^2 = Comunalidad.

5.3.- Resultados

En la Tabla 5.1 se presentan los resultados descriptivos de los ítems de la escala HoNOS y sus características de consistencia interna. Cinco de los doce ítems presentan efectos suelo notorios con porcentajes de respuesta superiores al 50% en la respuesta de ausencia del síntoma (valor 0) y, por tanto, con asimetrías positivas superiores al valor unidad. El ítem 9, por el contrario, es el único que presenta asimetría negativa, pero con un valor dentro de la normalidad ($As=-.19$).

La correlación promedio del ítem con el total de la escala es moderada ($r=.34$), con valores que oscilan entre .17 y .65. Por su parte, la consistencia interna para el total de la escala es alta ($\alpha=.70$), y la retirada de ninguno de sus componentes logra mejorar sustantivamente este valor de fiabilidad. No obstante, si se atiende a la configuración dimensional propuesta por Wing y colaboradores (1998) los valores de consistencia interna resultan bajos (valores α entre .10 y .45), exceptuando la dimensión de problemas sociales que ofrece un coeficiente ($\alpha=.68$) próximo a la consideración de alto.

El análisis de la matriz de correlaciones realizada con el programa FACTOR presenta un valor del coeficiente de Mardia de 65.12, indicando la presencia de asimetría multivariada, por lo que se sugiere la utilización de matrices policóricas en el análisis de la estructura factorial de los ítems. Asimismo, la prueba de Análisis Paralelos (Parallel Analysis) y el Método Hull utilizados para la extracción de factores sugieren la retención de un único factor. En la Tabla 2.1 (pg. 48) se han presentado las distintas estructuras factoriales sugeridas en la literatura para el HoNOS. De la propuesta original de cuatro factores de Wing et al (1998) a otras alternativas también de cuatro factores (Preston, 2000, McClelland et al., 2000; Newnhan et al., 2009) o de cinco factores (Trauer, 1999), e incluso de un único factor con restricción de ítems a un número de seis (Lovaglio & Monzani, 2012) o de cuatro elementos (Speak & Muncer, 2015). Las diferentes propuestas dimensionales han sido probadas mediante técnicas de análisis factorial confirmatorio en la configuración propuesta por cada autor, así como en el caso de dichas configuraciones bajo el supuesto de un modelo bifactorial de un factor general y las correspondientes dimensiones propuestas en cada modelo. En la Tabla 5.2 se presentan los índices de ajuste de cada uno de los modelos sometidos a prueba. El modelo que ha mostrado unos peores índices de ajuste ha sido el propuesto originalmente por Wing, mejorando los resultados cuando los cuatro factores se han correlacionado. Teniendo en cuenta los índices robustos y los cálculos de la invarianza entre modelos,

Tabla 5.2.- Índices de los Análisis Factoriales Confirmatorios de los distintos modelos estructurales (índices robustos)

Modelo	χ^2	gl	χ^2 / gl	AIC	GFI	NNFI	CFI	SMSR	RMSEA	LI	LS
Un factor	126.95	54	2.35	18.96	.96	.85	.88	.069	.069	.054	.085
Wing (4-factores)	382.96	54	7.09	274.96	.83	.32	.45	.150	.148	.133	.161
Wing (4-factores correlacionados)	100.96	48	2.10	4.96	.97	.88	.91	.064	.063	.045	.080
Trauer (5-factores cor.)	79.68	43	1.85	-6.32	.98	.91	.94	.052	.055	.036	.074
McClelland (4-factores cor)	93.12	48	1.94	28.13	.93	.80	.85	.063	.058	.040	.075
Newnhan (4-factores cor.)	82.41	45	1.83	-7.59	.98	.91	.94	.061	.054	.035	.073
Lovaglio	40.72	9	4.52	22.73	.98	.79	.88	.064	.112	.078	.148
Speak (4 factores cor)	96.68	47	2.05	2.68	.97	.88	.92	.063	.061	.044	.079
Bifactorial – Wing	50.43	36	1.40	-21.56	.99	.95	.97	.045	.038	.000	.061
Bifactorial - McClelland	60.14	36	1.67	-11.86	.95	.85	.92	.048	.049	.026	.070
Bifactorial – Newnham	44.54	36	1.24	-27.45	.99	.97	.98	.041	.029	.000	.054
Bifactorial – Speak	52.31	35	1.49	-17.69	.99	.95	.97	.046	.042	.013	.064

Nota. χ^2 : Prueba Ji cuadrado de Bondad de Ajuste; gl: Grados de Libertad; AIC: Criterio de Información de Akaike; GFI: Índice de Bondad Global; NNFI: Índice de Ajuste No Normado; CFI: Índice de Ajuste Comparativo; SMSR: Residuo Cuadrático Medio Estandarizado; RMSEA: Error Cuadrático Medio de Aproximación; LI - LS: Límite de Confianza Inferior - Superior del 90% del RMSEA

podemos observar cómo las estructuras bifactoriales propuestas por los distintos autores ajustan mejor. El modelo bifactorial con cuatro factores correlacionados de Wing et al (1998) [$\chi^2(36)=50.42$, $p<.001$, $AIC=-21.56$, $CFI=.97$, $NNFI=.95$, $RMSEA=.038$ (.000-.061)], el bifactorial de cuatro factores correlacionados de McClelland et al. (2000) [$\chi^2(36)=60.14$, $p<.001$, $AIC=-11.86$, $CFI=.92$, $NNFI=.82$, $RMSEA=.049$ (.026-.070)], el bifactorial de cuatro factores correlacionados de Newnham et al (2009) [$\chi^2(36)=44.54$, $p<.001$, $AIC=-27.45$, $CFI=.98$, $NNFI=.97$, $RMSEA=.029$ (.000-.054)] y el modelo bifactorial de cuatro factores correlacionados de Speak et al. (2016) [$\chi^2(35)=52.31$, $p<.001$, $AIC=-17.69$, $CFI=.97$, $NNFI=.95$, $RMSEA=.042$ (.013-.064)], obteniendo unos índices adecuados.

Los datos de la validez concurrente de la escala HoNOS se presentan en la Tabla 5.3. Las distintas dimensiones del HoNOS correlacionan significativamente y de forma congruente con los índices de valoración clínica: de forma negativa con el EEAG (una mayor gravedad en las escalas HoNOS se asocian a una peor funcionalidad reflejada por bajas puntuaciones en la EEAG) y de forma positiva con el ICG (una mayor puntuación en ambos instrumentos indica una mayor gravedad o severidad sintomática). Por su parte, los valores

Tabla 5.3.- Validez concurrente

	HoNOS			
	Problemas conductuales	Deterioro	Problemas clínicos	Problemas sociales
EEAG – Evaluación Actividad Global	-.276**	-.305**	-.453**	-.484**
ICG – Impresión Clínica Global	.270**	.332**	.436**	.482**
EuroQol – Calidad de Vida	-.044	-.019	-.300**	-.123*
HoNOS				
Problemas conductuales	1			
Deterioro	.126*	1		
Problemas clínicos	.271**	.100	1	
Problemas sociales	.362**	.362**	.368**	1

Nota. * $p < .05$ y ** $p < .001$

de correlación de la calidad de vida percibida por el participante respecto a las valoraciones HoNOS del clínico, sólo han resultado estadísticamente significativas para las dimensiones de problemas clínicos ($r = -.30$) y problemas sociales ($r = -.12$). Finalmente, la asociación entre las distintas dimensiones del HoNOS han resultado significativas y en sentido positivo para las diferentes combinaciones, salvo en el caso de la asociación entre los problemas clínicos y el deterioro ($r = .10$) que ha resultado estadísticamente no significativa.

Tabla 5.4.- Diferencias por tipo de centro asistencial en las puntuaciones HoNOS

	CSM - a (n= 164)		HD - b (n= 74)		TAC - c (n= 22)		HP - d (n= 21)		Prueba estadística de contraste de medias		
	M	DE	M	DE	M	DE	M	DE	F	p	Post-Hoc
Probl. conductuales	0.84	1.36	0.82	1.31	0.91	1.85	2.19	2.06	5.606	.001	ad,bd,cd
Deterioro	2.16	1.48	2.39	1.45	1.91	1.60	3.57	1.50	6.201	<.001	ad,bd,cd
Problemas clínicos	3.14	2.22	3.58	2.71	3.27	1.98	2.67	1.62	1.072	.361	
Problemas sociales	4.03	2.86	4.96	2.56	5.36	3.95	9.90	2.26	26.962	<.001	ad,bd,cd
Total	10.18	5.81	11.76	5.47	11.45	5.92	18.33	4.16	13.296	<.001	ad,bd,cd

Nota. CSM = Centro de Salud Mental, HD = Hospital de Día, TAC = Tratamiento Asertivo comunitario, HP = Hospitalización Psiquiátrica, M = Media, DE = Desviación Estándar, Post-Hoc: letras iguales expresan diferencias estadísticamente significativas entre los grupos que representan

Por último, en la Tabla 5.4 están los resultados del ANOVA para el contraste de diferencias en las puntuaciones del HoNOS atendiendo al dispositivo asistencial donde el participante recibe atención. Para la puntuación total del instrumento los valores más altos se encuentran entre los participantes asistidos en Hospital Psiquiátrico ($M=18.33$), cuya mayor expresión sintomática es estadísticamente significativa respecto a los otros tres tipos de recursos asistenciales ($M_{\text{Hospital de Día}}=11.76$; $M_{\text{Tratamiento Asertivo Comunitario}}=11.45$; y $M_{\text{Centros de Salud Mental}}=10.18$), los cuales no difieren entre sí. Un similar patrón de diferencias se encuentra para las tres dimensiones del HoNOS que han resultado estadísticamente significativas (problemas conductuales, deterioro y problemas sociales).

Capítulo 6

Estudio 2: Nuevas Evidencias de Adecuación Psicométrica de la Versión Española del HoNOS

6.1.- Objetivo

Ofrecer nuevas evidencias de adecuación psicométrica de la escala HoNOS en población de Personas con Enfermedad Mental Grave (PEMG) atendidas en el servicio de la Red de Salud Mental de Bizkaia (RSMB) (n= 3108). En este caso, se trata de reunir las evidencias de fiabilidad (consistencia interna y fiabilidad temporal -test retest-) y validez (constructo, convergente, discriminante y predictiva) de la escala HoNOS obtenidas tanto a través de la Teoría Clásica de los Test (TCT) como de la Teoría de Respuesta al Ítem (TRI).

6.2.- Método

6.2.1.- Muestra

La muestra total está compuesta por 3715 episodios, 1446 de ellos son de mujeres (37.5%) y 2403 de hombres (62.3%). La media de edad es 50,21 años para las mujeres (DT 12.49) y 45.69 años para los hombres (DT 11.63). La muestra se ha clasificado en base al tiempo de recorrido en la RSMB, incidentes si la duración ha sido menor o igual a un año o prevalentes si la duración del tratamiento ha sido mayor de un año. Los diagnósticos se han clasificado en base a los criterios para el Trastorno Mental Grave (TMG) el cual se define como “un grupo de personas heterogéneas, que sufren trastornos psiquiátricos graves que cursan con alteraciones mentales de duración prolongada, que conllevan un grado variable de

discapacidad y de disfunción social, y que han de ser atendidas mediante diversos recursos sociosanitarios de la red de atención psiquiátrica y social”. Se contemplan tres dimensiones: diagnóstico, cronicidad y discapacidad. (Ministerio de Sanidad y Política Social, 2009).

La dimensión de diagnóstico incluye los siguientes criterios diagnósticos de la Clasificación Internacional de Enfermedades (CIE-10) (Organización Mundial de la Salud, 2003): trastorno esquizofrénico (F20.x), trastorno esquizotípico (F21), trastorno delirante persistente (F22), trastorno delirante inducido (F24), trastorno esquizoafectivo (F25), otros trastornos psicóticos no orgánicos (F28 y F29), trastorno bipolar (F31.x), episodio depresivo grave con síntomas psicóticos (F32.3), trastorno depresivo grave recurrente (F33) y trastorno obsesivo compulsivo (F42). Las características de la muestra están recogidas en la Tabla 6.1.

Tabla 6.1.- Características sociodemográficas de la muestra

		Incidentes (n=526)	Prevalentes (n=3189)
Sexo [n(%)]	Hombres	338 (64.2)	1972 (61.8)
	Mujeres	183 (37.8)	1214 (38.2)
Edad [M(DE)]		48.65 (13.44)	47.38 (11.91)
Recurso terapéutico [n(%)]	Centro de salud mental -CSM-	106 (20.2%)	1450 (45.5%)
	Hospital	322 (61.2%)	618 (19.4%)
	Centro de día	41 (7.8%)	792 (24.8%)
	Tratamiento Asertivo Comunitario -TAC-	57 (10.8%)	329 (10.3%)
Diagnóstico [n(%)]	Esquizofrenia	194 (36.9%)	1699 (53.3%)
	Trast. Esquizotípico	3 (0.6%)	19 (0.6%)
	Ideas delirantes persistentes	23 (4.4%)	134 (4.2%)
	Trast. Psicóticos agudos	3 (0.6%)	106 (3.3%)
	Trast. Esquizoafectivos	24 (4.6%)	238 (7.5%)
	Otros Trast. psicóticos no orgánicos	-	18 (0.6%)
	Psicosis no orgánica sin especific.	14 (2.7%)	63 (2.0%)
	Trast. Bipolar	28 (5.3%)	283 (8.9%)
	Episodios depresivos graves	2 (0.4%)	34 (1.1%)
	Trast. Depresivo recurrente	1 (0.2%)	17 (0.5%)
	Trast. Obsesivo compulsivo	7 (1.3%)	49 (1.5%)
	Trast. Afectivos, ansiedad y neuróticos	5 (1%)	169 (5.3%)
	Otros	31 (5.9%)	353 (11.1%)
	Sin diagnóstico	191 (36.3%)	5 (0.2%)

6.2.2.- Instrumento

La escala HoNOS (Wing, Beevor y Curtis, 1996), adaptada al castellano por Uriarte (Uriarte et al, 1999) está constituida por un conjunto de escalas diseñadas para medir los diversos problemas físicos, personales y sociales asociados a la enfermedad mental. Los profesionales de la Salud Mental pueden hacer un uso clínico de él, de forma rutinaria y en un contexto clínico. Permite un uso administrativo, generando un eventual sistema nacional de recogida de datos con información clínica relevante para la medición de la evolución de los pacientes y la planificación de decisiones en salud pública. Contiene 12 ítems y se puntúan de 0 a 4. La puntuación total representa la severidad global, siendo la máxima puntuación posible de 48; a mayor puntuación, mayor gravedad. Se evalúa al inicio y al final de la intervención y en general se consideran las dos semanas previas, aunque pueden establecerse otros periodos según el contexto.

6.2.3.- Procedimiento

El programa de Trastorno Mental Grave de la Red de Salud Mental de Bizkaia tiene implementado un protocolo de valoración del estado de gravedad y evolución en el programa de los usuarios que acceden al mismo. La valoración de cada sujeto en el momento de evaluación que le corresponde es realizada por el equipo que le asiste a través de los instrumentos HoNOS, ICG y EEAG. Para el presente estudio se tomó como base de análisis las primeras valoraciones o valoraciones basales de HoNOS recogidas para cada nuevo episodio entre los años 2013-2016 hasta noviembre de 2020.

6.2.4.- Aspectos éticos

La información de las personas participantes en el estudio se trató de forma anónima según el procedimiento de la Ley Orgánica 15/1999 de 13 de diciembre de protección de datos de carácter personal. Este proyecto de investigación respetó los principios fundamentales establecidos en la Declaración de Helsinki, así como los requisitos establecidos por la legislación vigente en el ámbito de la investigación médica. Con el objetivo de preservar la integridad y confidencialidad de los datos del estudio, estos fueron extraídos con un sistema de anonimización que garantizaba el cumplimiento de la Ley Oficial de Protección de Datos (LOPD) y que permitía garantizar el anonimato de la información clínica, y así se evitase vinculación alguna con los datos de la historia clínica del participante. Dicho procedimiento fue realizado por el equipo técnico de la RSMB, ofreciendo al equipo

investigador los datos desagregados y anonimizados, impidiendo cualquier tipo de identificación personal. En el caso de necesitar información complementaria, cada sujeto de análisis estaba asociado a un código numérico que se vinculaba con su historia clínica; no obstante, la búsqueda de información, cumplimentación del registro y remisión anonimizada al equipo de investigación siempre se hacía desde el equipo técnico de la RSMB.

En todo momento el procedimiento de evaluación sigue los estándares de comportamiento ético en la investigación y cuenta con la aprobación del CEIC (acta 8/2012 10 de octubre de 2012) (Comité Ético de Investigación Clínica de Euskadi - Osakidetza).

6.2.5.- Análisis de datos

Para el análisis de los ítems de HoNOS se utilizaron la frecuencia (n), porcentaje, media (M), desviación estándar (DE), asimetría (As), correlación entre el ítem con el resto de la escala (r), la comunalidad (h^2), Alfa de Cronbach (α) si se elimina elemento y Alfa de Cronbach para el total y las dimensiones. Para el análisis de asociación entre variables se utilizó el coeficiente de correlación de Pearson (r). Se calcularon los índices de Mokken (H, H_i , H_{ij}) junto con los análisis de la monotonía manifiesta y las intersecciones para los ítems y por escalas, a través del paquete ‘mokken’ en R. Se realizó el análisis de las curvas características de los ítems y se calcularon los parámetros del modelo de respuesta graduada de Samejima a través del paquete ‘ltm’ en R. El análisis factorial exploratorio se realizó a través del análisis paralelo de Horn, mientras que los análisis factoriales confirmatorios de realizaron a través del paquete ‘lavan’ en R. Dada la asimetría encontrada ya que se tratan de ítems tipo likert se utilizó el método robusto de DWLS, es decir, WLSMV.

6.3.- Resultados

En la Tabla 6.2 se recogen los resultados descriptivos de los ítems de la escala HoNOS y las características de consistencia interna. Si analizamos los estadísticos descriptivos para el total de la muestra y las medidas de tendencia central de los ítems vemos que se produce un efecto suelo en los ítems 1, 2, 3 y 11, ya que más del 50% de las puntuaciones son 0 (ausencia de síntoma) (se encuentran en la cola inferior). Debido a esto, las asimetrías en los tres primeros ítems están especialmente marcadas superiores al valor unidad (1.08, 3.95 y 1.33 respectivamente). Por el contrario, los ítems 9 y 10 presentan asimetría negativa con valores dentro de la normalidad (-0,46 y -0,19, respectivamente).

Tabla 6.2.- Estadísticos descriptivos y medidas de tendencia central de los ítems y dimensiones del HoNOS (n=3715)

Ítems	Puntuaciones					M	DE	As	H _i	r	α_i	h ²
	0	1	2	3	4							
1	52.9	19.4	16.4	7.8	3.5	0.90	1.14	1.08	0.28	.49	.697	0.99
2	89.6	5.2	2.7	1.5	1.0	0.19	0.64	3.95	0.24	.26	.726	0.21
3	64.6	10.6	11.6	8.9	4.3	0.78	1.21	1.33	0.16	.26	.727	0.16
4	25.3	29.8	29.3	13.2	2.3	1.38	1.07	0.32	0.22	.38	.712	0.21
5	38.8	18.3	24.5	15.0	3.3	1.26	1.21	0.45	0.08	.13	.742	0.06
6	30.7	18.8	23.5	18.5	8.5	1.55	1.32	0.28	0.17	.29	.724	0.21
7	42.5	25.1	22.1	8.6	1.6	1.02	1.07	0.72	0.14	.23	.728	0.42
8	35.0	13.3	25.7	20.8	5.2	1.48	1.29	0.21	0.18	.31	.721	0.25
9	9.9	13.8	29.1	36.8	10.3	2.24	1.12	-0.46	0.31	.51	.696	0.66
10	18.8	15.5	30.3	28.3	7.0	1.89	1.21	-0.19	0.35	.61	.680	0.57
11	52.2	12.0	12.4	8.2	15.1	1.22	1.51	0.82	0.25	.44	.702	0.50
12	38.9	13.4	19.1	15.9	12.8	1.50	1.45	0.39	0.28	.51	.692	0.49
						M	DE	As	H		α_r	
D1						1.86	2.14	1.25	0.30	.56	.460	.34
D2						2.63	1.73	0.37	0.15	.42	.245	.19
D3						4.05	2.46	0.38	0.17	.53	.370	.25
D4						6.85	3.98	0.26	0.48	.74	.734	.22
Total						15.40	7.27	0.31	0.22		.731	

Nota. 0=Sin problema, 1=problema menor que no requiere intervención, 2=problema leve pero claramente presente, 3=problema de moderada gravedad o 4=problema grave o muy grave, M=media, DE=Desviación estándar, As=Asimetría, H_i= índice H para los ítems, H= índice H para el total de la escala y subescalas, r= correlación con la escala, α_i =Alfa de Cronbach si se elimina elemento y α_r para escala total y subescalas, h²= Comunalidad.

En cuanto a los índices H_i todos son inferiores a .30 salvo en los ítems 9 y 10 (.31 y .35). En referencia a la correlación de los ítems con la escala 5 de los 12 ítems se correlaciona con valores mayores de .40 (ítems 1, 9, 11 y 12). Del resto de ítems solo dos tienen valores mayores al mínimo exigible (.30; ítems 4 y 8), mientras que los otros cinco ítems tienen valores por debajo de .30 (ítems 2, 3, 5, 6, 7). En cuanto a los índices de fiabilidad Alpha de Cronbach, el valor más bajo es de .68 para el ítem 10, y .74 para el ítem 5. Los ítems que más contribuyen a la fiabilidad de la escala son 1, 9, 10 y 12. En 11 de los 12 ítems, si estos se eliminaran la fiabilidad de la escala bajaría, y tan solo en el caso del ítem 5 la fiabilidad total aumentaría en caso de eliminarlo, aunque no resultaría en un aumento sustancial (de .73 a .74).

A lo largo de las diversas puntuaciones ciertos ítems presentan claras asimetrías, como se puede observar en la Tabla 6.2, acumulando un gran porcentaje de respuestas en los valores bajos, destacando el ítem 2 que acumula un 89.6% de respuestas en el valor '0' provocando un efecto suelo en la respuesta de los ítems. Estos casos se dan principalmente en los tres primeros ítems, quienes conforman el primer factor, mostrando medias cercanas a '0' y asimetrías de más de '1'. Se muestra como el factor 1 presenta a su vez mayor a simetría que el resto. No se satisface el criterio de Mokken al encontrar coeficientes de escalabilidad (H_{ij}) menores a 0 en la matriz. Los ítems 5 y 7 muestran, por su parte, los índices de escalabilidad H_i más bajos, lo que concuerda con menores correlaciones y con los alphas más elevados en caso de su eliminación. A excepción de los ítems 9 y 10, ninguno de los restantes alcanza el valor mínimo de 0.30. Mostrando índices H bajos en el segundo y tercer factor, situándose el primero en el límite, y siendo el factor 4 el único que muestra una escalabilidad media, a su vez, la escalabilidad total sigue sin alcanzar el valor mínimo. Al realizar el análisis de la monotonía manifiesta y de las intersecciones se encuentran violaciones significativas en ambos casos. Dado los resultados encontrados no se puede asumir que nos encontremos ante un modelo de homogeneidad monótona. En el caso de las comunilidades, se puede observar ítems con valores altos, como son el 1 o los cuatro últimos, y como el ítem 5 muestra una comunilidad de prácticamente cero.

En la Tabla 6.3 vemos los valores para la muestra de pacientes que llevan menos de 1 año en la RSMB (incidentes, $n=526$) y para los pacientes que llevan más de 1 año en la Red (prevalentes, $n=3189$). Las puntuaciones medias obtenidas son mayores en la muestra de incidentes para todos los ítems salvo en el ítem 3 (Consumo Problemático de Alcohol o Drogas) y el ítem 7 (Problemas en relación con el humor depresivo). Los episodios incidentes presentan asimetrías marcadas en varios de los ítems, con valores superiores a 1 o rondando el 1, junto con medias cercanas a 0 o 1 en varios casos, como se muestra en la Tabla 7.3. Se observa cómo esta tendencia a valores bajos también se encuentra en los episodios prevalentes en los primeros ítems. Al realizar la comparación de medias de los ítems entre las muestras se encuentran diferencias estadísticamente significativas para 8 de 12 ítems, con tamaños del efecto bajos en su mayoría, a excepción de los ítems 11 y 12. También se encuentran diferencias estadísticamente significativas entre los factores, resultando el cuarto factor Social quien muestre un mayor tamaño del efecto.

Tabla 6.3.-Datos descriptivos y t de Student para la diferencia de medias entre Incidentes y Prevalentes

ítem	Rango	Incidentes (n=526)				Prevalentes (n=3189)				Pruebas estadísticas		
		M	SD	As	H _i	M	SD	As	H _i	t	p	d
1.	0-4	1.13	1.27	0.77	0.24	0.86	1.12	1.13	0.28	-4.67	.001	.236
2.	0-4	0.21	0.68	3.90	0.17	0.19	0.64	3.97	0.25	-0.58	.557	.031
3.	0-4	0.73	1.23	1.50	0.13	0.78	1.20	1.30	0.17	0.96	.334	.042
4.	0-4	1.60	1.10	0.11	0.18	1.34	1.06	0.35	0.22	5.28	.001	.244
5.	0-4	1.33	1.22	0.47	0.11	1.25	1.21	0.45	0.07	1.39	.163	.067
6.	0-4	1.83	1.41	0.08	0.11	1.51	1.30	0.31	0.17	4.81	.001	.243
7.	0-4	0.88	1.05	0.99	0.09	1.04	1.07	0.67	0.15	-3.13	.002	.150
8.	0-4	1.56	1.31	0.13	0.16	1.47	1.29	0.22	0.18	1.56	.118	.069
9.	0-4	2.40	1.15	-0.60	0.29	2.21	1.12	-0.44	0.31	3.59	.001	.169
10.	0-4	2.27	1.15	-0.59	0.30	1.83	1.21	-0.13	0.36	8.13	.001	.366
11.	0-4	2.14	1.78	-0.17	0.23	1.07	1.41	1.00	0.25	13.14	.001	.729
12.	0-4	2.11	1.50	-0.13	0.27	1.40	1.42	0.48	0.27	10.44	.001	.496
Subescalas												
Conducta	0-10	1.72	1.83	1.15	0.24	1.52	1.77	1.27	0.31	2.35	.019	.112
Deterioro	0-10	3.66	2.22	0.16	0.18	3.23	2.14	0.40	0.14	4.25	.001	.200
Síntomas	0-10	3.56	2.00	0.33	0.11	3.34	2.06	0.39	0.19	2.19	.028	.107
Social	0-10	5.57	2.57	-0.39	0.43	4.07	2.41	0.35	0.47	12.57	.001	.616
Total	0-10	3.79	1.49	-0.06	0.19	3.11	1.50	0.38	0.22	9.59	.001	.454

Notes. M=Media, DE= Desviación Estándar; As=Asimetría; H_i= Índice de Homogeneidad H para los ítems; H= Índice de Homogeneidad H para la escala y subescalas; t= Prueba t de Student; p= valor de significación; d= Tamaño del efecto de Hedges

En cuanto a los índices H_i, en ambas muestras el ítem 10 es el único en alcanzar el valor mínimo de .30, y en el caso de los prevalentes también el ítem 9, se encuentran coeficientes de escalabilidad H_{ij} menores de 0, no satisfaciendo el criterio de Mokken. A su vez, el factor Social muestra un valor medio de escalabilidad, sumándose en los episodios prevalentes el primer factor de problemas de conducta. En ninguna de las dos muestras el total de los episodios presenta un índice H que alcance el mínimo establecido. Al realizar los análisis de monotonía manifiesta e intersecciones se encuentran violaciones significativas tanto en los episodios prevalentes como en los incidentes, debido a esta serie de resultados no se puede

afirmar que nos encontremos ante modelos de homogeneidad monótona, ni para los episodios prevalentes ni para los incidentes.

Tras los análisis de Mokken se decidió realizar los análisis de partición en escalas de la muestra total y las dos submuestras (Incidentes y Prevalentes). Estos análisis reportan la conformación de dos escalas en el caso del total de casos y la muestra de prevalentes, no situando los ítems 3, 5 y 6 en ninguna escala, y en el caso de la muestra de incidentes se reportan tres escalas no situando en ninguna a los ítems 4, 5, 6 y 8. Al realizar los análisis de Mokken por escalas se decidió utilizar la estructura planteada principalmente por los autores a la recomendada por el análisis de partición encontrado.

Tabla 6.4.- Escalabilidad de Mokken (Hi) por escala y Samejima's graded response model parameters

Item	Total de casos			Casos Incidentes			Casos Prevalentes		
	Hi	a	Item Info	Hi	a	Item Info	Hi	a	Item Info
Conducta									
1	.35	1.19	2.63	.30	0.92	1.91	.36	1.23	2.74
2	.22	0.82	1.47	.15	0.54	0.79	.23	0.89	1.63
3	.29	0.57	1.00	.24	0.36	0.49	.31	0.62	1.13
Deterioro									
4	.15	0.99	2.59	.18	0.84	2.16	.14	0.99	2.60
5	.15	0.29	0.43	.18	0.43	0.82	.14	0.27	0.37
Síntomas									
6	.11	0.72	1.52	.08	0.58	1.11	.12	0.73	1.56
7	.21	0.46	0.91	.13	0.28	0.38	.23	0.54	1.17
8	.21	0.70	1.50	.13	0.78	1.69	.22	0.70	1.51
Social									
9	.42	1.53	4.14	.37	1.58	4.17	.42	1.53	4.18
10	.50	2.26	6.47	.42	1.76	4.68	.51	2.30	6.66
11	.47	1.30	2.21	.42	1.13	1.59	.46	1.28	2.29
12	.51	1.60	3.31	.50	1.57	3.16	.50	1.55	3.23

Notas:

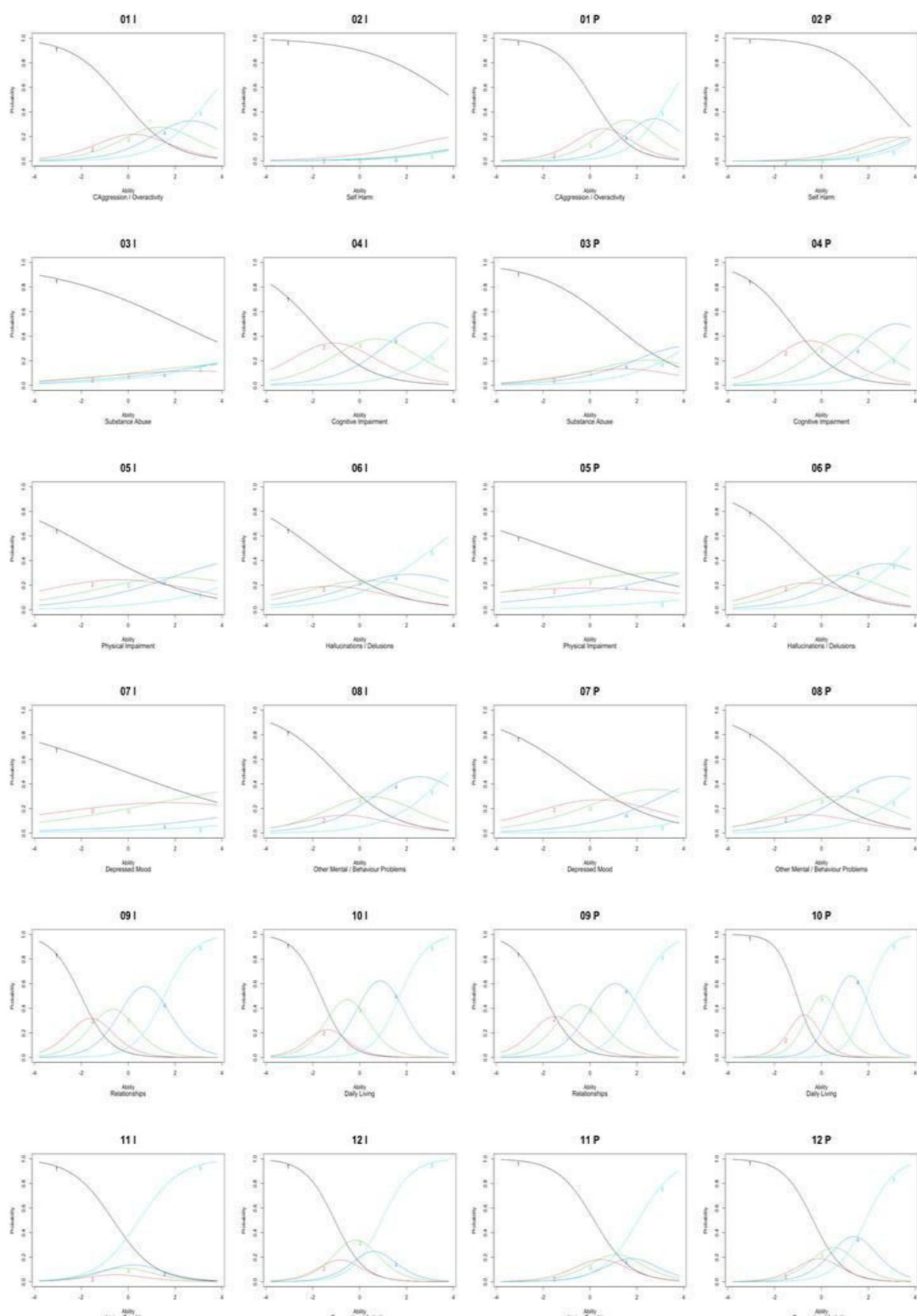


Figura 6.1.- Curvas características de los ítems en muestra incidente - prevalente

En la muestra total la cuarta escala (Social) muestra los mejores resultados, mientras las restantes presentan valores Hi bajos, y múltiples violaciones significativas, como se observa en la Tabla 6.4. En la escala Social se encuentran coeficientes de escalabilidad Hij superiores a 0, junto con índices de escalabilidad Hi bajos o moderados en los cuatro ítems y una escalabilidad H del factor moderada, cumpliendo con todos los criterios de Mokken. No encontrándose violaciones significativas en el análisis de la monotonía manifiesta, aunque sí intersecciones entre los cuatro ítems. Los resultados encontrados a través de la muestra de prevalentes son similares a la del total de los episodios. En las tres primeras escalas se encuentra Hi que no alcanzan el mínimo, además de múltiples violaciones significativas en los análisis de la monotonía manifiesta y las intersecciones. La cuarta escala de nuevo reporta los mejores valores, pero al igual que en el caso anterior se encuentran violaciones significativas en el análisis de las intersecciones. En el caso de la muestra de incidentes la escala cuatro es la única en reportar índices Hi que alcancen el mínimo, siendo en su caso también superior a .30 el valor H. Aunque encontrándose violaciones significativas tanto en el análisis de la monotonía manifiesta como en el de las intersecciones.

A través de los parámetros del modelo de Samejima se observa cómo tanto en el total como en los casos prevalentes los ítems que mejor discriminan son los concernientes a la escala 4 junto al ítem 1, mientras que el ítem 5 debería considerarse para su eliminación. En el caso de la muestra de incidentes los ítems de la subescala cuatro (9, 10, 11 y 12) también serían los que mejor ajuste mostrarían, aunque en este caso el ítem que se recomendaría para ser eliminado sería el 7.

6.5.- Índices de Ajuste de los modelos para la muestra total y las submuestras (Incidentes y Prevalentes)

Modelo	Muestra	χ^2	gl	χ^2 / gl	GFI	AGFI	CFI	NNFI	SRMR	RMSEA	LI	LS
4FR	Total	1869.51	48	38.95	.980	.952	.870	.822	.074	.101	.097	.105
	Incidente	328.70	48	6.85	.969	.926	.841	.744	.087	.106	.095	.116
	Prevalente	1564.41	48	32.59	.981	.954	.875	.828	.073	.100	.095	.104
MA	Total	2090.39	50	41.81	.977	.948	.855	.808	.079	.105	.101	.109
	Incidente	351.98	50	7.04	.966	.924	.799	.735	.091	.107	.097	.118
	Prevalente	1761.70	50	35.23	.978	.950	.858	.813	.078	.104	.100	.108

Notas.- 4FR = Modelo de 4 factores correlacionados, MA = Modelo anidado.

La Tabla 6.5 presenta los resultados en los índices de ajuste para las distintas muestras en los dos modelos planteados. Por un lado, se encuentra el modelo en el que existen 4 factores correlacionados entre sí que cargan a los ítems. Por otro lado, se encuentra el modelo anidado en el cual un factor que representaría carga en los cuatro factores, que no correlacionan, y que a su vez cargan en los ítems. Se puede observar en todos los casos Ji cuadrados elevados, lo que genera que al dividirlos entre los grados de libertad supere los valores óptimos por debajo de 3. En ambos modelos se observa como los episodios prevalentes muestran índices de ajuste mejores a los totales e incidentes. A pesar de su mejor actuación solo muestran un buen ajuste en el GFI y AGFI (que no son índices robustos) y un ajuste aceptable en SRMR. En líneas generales no existen prácticamente diferencias de ajuste entre los modelos, no se encontraría ajuste de los datos a los modelos, a pesar de los ajustes encontrados no robustos y el pobre ajuste que se presentaría en algunos indicadores del RMSEA.

Al encontrar un número de escalas diferentes en la partición de Mokken al originalmente planteados por los autores, se decidió realizar análisis paralelos de Horn para conocer la posibilidad de un número distinto de factores. Se encontraron para las tres muestras (total, prevalentes e incidentes) soluciones factoriales de 5 factores, lo cual difiere tanto con la estructura original como por los resultados de la partición en escalas de Mokken.

Capítulo 7

Estudio 3: Baremación de la escala HoNOS en población bizkaína

7.1.- Objetivo

Ofrecer baremos de las puntuaciones de HoNOS para la clasificación y evaluación de resultados en PEMG de la RSMB, y puntos de corte y valores de referencia de cambio clínico a través de la Puntuación Fiable de Cambio atendiendo a las variables de segmentación sexo y servicio de atención clínica (de forma independiente o combinada).

7.2.- Justificación

El objetivo principal de cualquier servicio de salud mental es mejorar el funcionamiento y la salud de las personas con trastornos mentales. Cada vez se hace más hincapié en la medición directa de los resultados de salud a través de medidas estandarizadas que tengan fiabilidad y validez. A nivel internacional, existe un fuerte impulso para el seguimiento de los resultados en salud mental.

La evaluación de resultados en el ámbito clínico se nutre de la aplicación de los hallazgos que surgen de la investigación, tanto en el desarrollo de instrumentos de medida fiables y válidos como en la aplicación de estadísticos que nos permitan interpretar los resultados obtenidos. Estos últimos nos permiten mostrar cuándo los cambios en las puntuaciones tienen

implicaciones que se plasman a nivel práctico y son aplicables en el mundo del sujeto (Pirkis et al., 2005). Si se consigue demostrar que la puntuación de un sujeto se ha trasladado desde el rango de la población “no funcional” a la “funcional” durante el tiempo de aplicación del tratamiento, se habrá producido un cambio clínicamente significativo (Iraurgi, 2009 y 2010).

Jacobson, Follette y Revenstorf (1984) y Jacobson y Truax (1991) desarrollaron una serie de métodos para identificar el cambio clínicamente significativo. Entendiendo éste como el grado en que una terapia o intervención desplaza a un sujeto fuera del rango de lo que se comprende como “población disfuncional” o lo introduce en el rango de la “población funcional”, e infiriendo de esta forma que se ha producido un cambio clínicamente significativo. Además de este criterio para entender el cambio clínico, también se desarrolló el Índice de Cambio Fiable (RCI – en su acrónimo en inglés: Reliability Change Index), un índice estadístico que valora la magnitud del cambio logrado, teniendo en cuenta la precisión del instrumento de medida empleado (Jacobson y Truax, 1991).

Las políticas de salud tienen como objetivo final establecer normas nacionales de resultados que permitan evaluar y valorar los distintos sectores que conforman los servicios de salud mental. Una vez establecidos estos estándares, la evaluación rutinaria de resultados formará parte integral de una buena atención clínica, ya que la información obtenida identificará el resultado del tratamiento y podrá ser de apoyo para informar y mejorar la prestación en los servicios (Burgess., 2015).

7.3.- Método

Los datos de caracterización de la muestra, procedimiento, ética e instrumento utilizado ya han sido expuestos en el capítulo anterior (pg. 89 a 92), al tratarse de la misma base de datos. Por ello, no se desarrollarán en este espacio, pero sí se detallará la estrategia de análisis utilizada para la baremación y cálculo de índices de cambio.

7.3.1.-Análisis estadísticos

Se estimó la media (M), desviación estándar (DE) y los percentiles 95, 90, 85, 80, 75, 50, 25, 20, 15, 10 y 5. Estos se llevaron a cabo a través del método clásico o frecuentista para el total de la escala, así como para cada una de las cuatro subescalas (problemas de conducta, deterioro, problemas clínicos y problemas sociales). Por una parte, se calcularon en función

del sexo y centro asistencial y por otra, en función del tiempo que llevaban en el PTMG (incidentes -menos de un año- y prevalentes -más de un año-) y centro asistencial.

7.3.1.1.- Cambio Clínicamente Significativo e Índice de Cambio Fiable -RCI

La primera aproximación a una medida del cambio clínicamente significativo fue propuesta por Jacobson, Follette y Reventorf en 1984, y posteriormente revisada por Jacobson y Truax (1991) en el ya célebre artículo ‘Clinical significance: A statistical approach to defining meaningful change in psychotherapy research’. El método propuesto por Jacobson y Truax, (RCI) consiste en dos partes diferenciadas: la primera consiste en la definición de los criterios que expresan qué ha de considerarse como ‘cambio clínico’, y la segunda la propuesta de un índice estadístico que valore el grado de significación en la precisión del cambio obtenido. El propósito de esta propuesta era: a) establecer un consenso sobre la definición de ‘cambio clínicamente significativo’ que sirviera para ser aplicado a cualquier trastorno clínico, b) determinar un punto de encuentro entre los profesionales y los usuarios respecto a lo que supone un cambio clínico en las expectativas de resultado en psicoterapia, y c) ofrecer un método preciso para la clasificación del resultado terapéutico alcanzado por los receptores de la intervención (mejoría, sin cambios, empeoramiento, etc.) en base a los criterios definidos como clínicamente significativos.

7.3.1.1.1.- Criterios para la definición del cambio clínico

De manera general, se puede plantear que la significación clínica de un tratamiento o intervención viene determinada por su capacidad para cumplir ciertas normas sobre su eficacia-efectividad establecidas por los protagonistas (pacientes, sanitarios, investigadores, ...) del tratamiento. No existe un consenso al respecto, dado que las voces son múltiples; pero para su utilización práctica, y con aceptable aprobación por las distintas partes, se han aceptado algunas como las siguientes: a) un alto porcentaje de personas tratadas que hayan mejorado con el tratamiento; b) un cambio en el estado de salud que sea reconocible por las personas que están alrededor de la persona tratada (Kazdin, 1977; Wolf, 1978); c) la eliminación de los signos y síntomas que definen el problema de salud (Kazdin y Wilson, 1978); d) alcanzar un estado de salud normativo al finalizar la terapia (Kendall y Norton-Ford, 1982; Nietzel y Trull, 1988); o e) cambios que reduzcan significativamente el riesgo de presentar problemas de salud (Mavissakalian, 1986).

Para Jacobson y Truax (1991), el punto de partida en la contextualización de la significación clínica tras una intervención terapéutica radica en el cambio hacia los valores

de normalidad dentro del área clínica sometida a intervención. Es decir, el cambio clínicamente significativo se produce cuando se evidencia la recuperación de un funcionamiento 'normal' de salud, entendiendo 'normal' desde una conceptualización de distribución poblacional del fenómeno. Es decir, estos autores plantean que el cambio clínicamente significativo se produce cuando el enfermo retorna a un funcionamiento normal, es decir, cuando puede ser considerado como parte de la población funcional y, por tanto, se precisarían datos de la distribución de puntuaciones respecto al indicador utilizado para establecer lo que es funcional y/o disfuncional.

Si la extracción muestral se obtiene del colectivo de personas sanas, es decir, aquellas que no presentan ningún trastorno o enfermedad -al menos en lo que respecta al trastorno de valoración-, hablaríamos de una 'población funcional'. Si, por el contrario, el estudio se realizase exclusivamente con personas que presentan el trastorno o enfermedad de interés nos hallaríamos con una 'población disfuncional'. En los estudios clínicos, como es obvio, se cuenta con muestras de enfermos -'poblaciones disfuncionales'- más o menos amplias y, en algunos casos, se cuenta con datos normativos. Lo que resulta menos habitual es disponer de poblaciones funcionales, y en ese caso el recurso es acudir a muestras normativas de población general -como se ha comentado previamente obtenidas a través de encuestas de salud o estudios comunitarios-, las cuales incluyen a personas que pueden presentar el trastorno o enfermedad de interés a nivel clínico o subclínico.

De este modo, teniendo en cuenta los distintos tipos de poblaciones descritas y las características de su distribución teórica, estos autores propusieron tres criterios, y sus correspondientes puntos de corte denominados como 'a', 'b' y 'c', en función de la combinación de ambas distribuciones (funcional y disfuncional) en el caso de contar con ambas. Consideremos que evaluamos sintomatología, como es el caso del HoNOS, y que contamos con una muestra suficiente y su distribución de puntuaciones. Tratándose de sintomatología, las poblaciones enfermas (disfuncionales) tenderán a reflejar altas puntuaciones como expresión de presencia de tales síntomas, de modo que una reducción de tales síntomas sería lo esperable como reflejo de una recuperación. Por tanto, obtener puntuaciones por debajo de la media sería un indicador de que se avanza hacia la funcionalidad o al menos de que uno se aleja de la funcionalidad. Ahora bien, cuánto ha de alejarse de dicha media. Lo convencional por criterios estadísticos-científicos, es equivocarnos pocas veces en nuestras decisiones, y por ello aceptamos no equivocarnos más de 5 veces de cada cien (5%), es decir un valor $\alpha < .05$, que son aproximadamente dos desviaciones típicas hacia la izquierda ($M_D - 2DE_D$), y a este punto es el que Jacobson y

Truax denominan 'a'. Por el contrario, en el caso de responder al indicador personas sanas, la mayor parte de éstas responderán con bajas puntuaciones, es decir, con baja sintomatología y, por tanto, la disfuncionalidad se situaría en las puntuaciones altas; esto es en la $M_F + 2DE_F$, señalándose de este modo el punto 'b'. Una vez que tenemos la forma de estas dos distribuciones, el punto 'c' sería el resultante de promediar las medias de ambas distribuciones (funcional y disfuncional) ponderadas por sus desviaciones estándar según la fórmula:

$$\text{Punto de Corte 'c'} = \frac{M_F \cdot DE_D + M_D \cdot DE_F}{DE_F + DE_D}$$

donde M_D y M_F son las medias de las muestras Funcional y Disfuncional, respectivamente, y DE_D y DE_F sus Desviaciones Estándar correspondientes.

En el caso que nos ocupa, HoNOS, solo contamos con datos muestrales de poblaciones clínicas, ya que es un instrumento que nunca se ha aplicado a población sana-normativa. No obstante, si se pueden diferenciar diferentes estados de gravedad en función de la progresión en el tratamiento de los usuarios. Es decir, los casos prevalentes (más de un año en el programa de trastornos mental grave -PTMG) marcados por una sintomatología más leve y con un mayor tratamiento a lo largo del tiempo, y los casos incidentes (menos de un año en el PTMG) con una sintomatología más aguda y un tiempo de tratamiento más corto. Estos dos grupos están diferenciados por la gravedad de los síntomas y el recurso asistencial en el cual se les presta apoyo. Es por esto por lo que nos ha resultado de gran interés establecer un punto de corte dentro del conjunto total de gravedad, que divida a estos dos subgrupos y permita realizar un tratamiento más concreto dependiendo del punto en el que se encuentren y el transcurso de su sintomatología (Jacobson y Truax, 1991).

7.3.1.1.2.- Criterios para la definición del cambio significativo

Jacobson y Truax (1991) proponen, como segunda condición para la consideración de cambio clínicamente significativo, la utilización de un índice que contemple la precisión de la medida; el aludido Índice de Cambio Fiable (RCI). Este índice fue propuesto inicialmente por Jacobson, Follette y Revenstorf (1984) y revisado posteriormente por Christensen y Mendoza (1986) y Jacobson y Revenstorf (1988). El algoritmo de cálculo se expresa del modo siguiente:

$$ICF_{JT} = \frac{X_{Post} - X_{Pre}}{EED} = \frac{X_{Post} - X_{Pre}}{\sqrt{2(DE_{Norma}\sqrt{1-r_{xx}})^2}}$$

donde

$$EED = \sqrt{2(EEM)^2} = EEM \sqrt{2}$$

y

$$EEM = DE_{Norma} \sqrt{1 - r_{xx}}$$

siendo X_{Post} la puntuación Post-tratamiento, X_{Pre} la puntuación Pre-tratamiento, EED el Error Estándar de las Diferencias entre las dos medidas, DE_{norma} la Desviación Estándar de la muestra normativa y EEM el Error Estándar de Medida del instrumento.

RCI se distribuye de forma ‘normal’, es decir, como una distribución de puntuaciones z . Si aceptamos los criterios de equivocarnos menos de un 5% de las veces, sabemos que corresponderían con valores $z_{.05} = 1.65$ en el caso de hipótesis unilaterales, o bien de $z_{.05/2} = 1.96$ en el caso de hipótesis bilaterales. Por tanto, si al realizar el cálculo de RCI obtenemos valores superiores al criterio adoptado, entonces nos encontraremos que el cambio observado ($X_{post} - X_{pre}$) es estadísticamente significativo y fiable. Es decir, de la formula inicial se puede derivar el cálculo de otro indicador: la Puntuación Mínima de Cambio (PMC), que se obtiene del siguiente modo:

$$RCI = (X_{post} - X_{pre}) / EED,$$

por tanto, y siendo

$$(X_{post} - X_{pre}) \rightarrow PMC$$

entonces

$$PMC = RCI * EED = 1.96 * EED$$

Como decimos, para que el RCI sea considerado estadísticamente significativo se toma como referencia la convención de un nivel de significación menor del 5% ($\alpha < .05$), tomando como referencia las puntuaciones z estandarizadas (Wilkinson, 1999). Debido a que solo contamos con la distribución de un grupo clínico y que el sentido de la funcionalidad se expresa en bajas puntuaciones, el 5% de los casos superiores se situarían por encima de un valor $z = 1.65$. Por lo tanto, teniendo en cuenta este criterio se calcula la Puntuación de la diferencia Mínima de Cambio (PMC), $PMC = z_{.05} * EED$, donde la puntuación estandarizada de $Z_{0.05}$ es 1.65.

Por último, para el cálculo del Punto de Corte (PC) se ha utilizado el mismo criterio que el mencionado para tomar como referencia el valor crítico del RCI, es decir, el punto de corte se establecería a partir del siguiente algoritmo: $PC_{\alpha.05} = M + 2DE * 1.65$. Sin embargo, debido a que el procedimiento propuesto por Jacobson y Truax de dos desviaciones estándar es un criterio estadístico muy restrictivo, también se han calculado los puntos de corte teniendo en cuenta otros dos criterios. El primero de ellos basado en criterios clínicos (Colton,

1995) que consistiría en restar una DE al valor de la media ($PC = M - 1DE$). Mientras que el segundo criterio vendría apoyado por evidencias de la ciencia básica (Miller, 1956; Norman, Sloan, & Wyrwich, 2003) que plantea la conveniencia de utilizar media desviación estándar ($PC = M - \frac{1}{2}DE$).

Ambos criterios (la PMC y los PC) han sido calculados al igual que los baremos teniendo en cuenta el sexo, tiempo en la Red y el centro asistencial. Además, la combinación de ambas condiciones permiten llevar a cabo una clasificación de los individuos (Jacobson & Truax, 1991) respecto a su consideración del cambio clínico en: ‘Recuperados’ (cuando supera ambos criterios), ‘Mejorados’ (supera la diferencia mínima de cambio, pero no se llega a alcanzar el punto de corte), ‘No cambio’ (no superan ninguno de los dos criterios, y/o ‘Deteriorados’ (el cambio es significativo y se produce en sentido contrario al de la funcionalidad) (Iraurgi, 2010).

7.4.- Resultados

En la Tabla 7.1 se presentan los datos descriptivos y el tamaño del efecto para la diferencia de medias entre el grupo de Incidentes y de Prevalentes, tanto para el total de la escala como para cada una de las cuatro subescalas del HoNOS. La prueba de t de Student es significativa para el total y las subescalas, obteniendo un tamaño del efecto moderado para el total y la subescala 4 (problemas sociales) (.45 y .62, respectivamente) y pequeño para las subescalas 1, 2 y 3 (.11, .20 y .11 respectivamente).

Tabla 7.1.- Datos descriptivos y t de Student para la diferencia de medias entre Incidentes y Prevalentes

	Total (n= 3715)				Incidentes (n= 526)				Prevalentes (n= 3189)				Prueba estadística		
	M	DE	EED	As	M	DE	EED	As	M	DE	EED	As	t	p	d
Escala total	15.40	7.27	0.12	0.32	18.18	7.17	0.31	-0.06	14.94	7.19	0.13	0.38	9.60	.001	.45
Problemas conductuales	1.86	2.14	0.03	1.25	2.07	2.20	0.10	1.15	1.83	2.13	0.04	1.27	2.36	.019	.11
Deterioro	2.63	1.73	0.03	0.37	2.93	1.78	0.08	0.16	2.58	1.71	0.03	0.40	4.25	.001	.20
Problemas clínicos	4.05	2.46	0.04	0.38	4.27	2.40	0.10	0.33	4.01	2.47	0.04	0.40	2.20	.028	.11
Problemas sociales	6.85	3.98	0.06	0.26	8.92	4.11	0.18	-0.34	6.51	3.85	0.07	0.35	13.16	.001	.62

Notas.- M: Media; DS: Desviación Estándar; EED: Error Estándar de las Diferencias; As: Asimetría; t: Prueba t de Student; p: probabilidad; d: Tamaño del efecto

Tabla 7.2.- Baremos para la escala total de HoNOS para la muestra total, muestra incidentes y prevalentes y para hombres y mujeres según el tipo de recurso asistencial

		TOTAL					HOMBRES					MUJERES				
		Total	CSM	HD	TAC	H	Total	CSM	HD	TAC	H	Total	CSM	HD	TAC	H
Total (n=3715)	n	3715	1556	833	386	940	2310	923	556	233	598	1397	632	277	151	337
	95	28.0	24.0	25.0	27.0	31.0	28.0	24.0	25.2	28.0	31.0	28.0	23.0	25.1	26.4	32.0
	90	25.0	21.0	22.0	25.0	29.0	25.0	21.0	23.0	25.6	29.0	25.0	21.0	21.0	24.0	30.0
	85	23.0	20.0	20.0	23.0	28.0	23.0	20.0	20.0	24.0	28.0	23.0	19.0	19.0	22.0	28.0
	80	22.0	18.0	18.0	21.0	27.0	22.0	18.0	19.0	22.0	26.0	22.0	18.0	18.0	20.6	27.0
	75	20.0	17.0	17.0	20.0	25.8	20.0	17.0	18.0	20.0	25.0	20.0	17.0	17.0	20.0	26.0
	50	15.0	12.0	14.0	15.0	21.0	15.0	12.0	14.0	16.0	21.0	15.0	11.0	14.0	15.0	22.0
	25	10.0	8.0	10.0	11.0	17.0	10.0	8.0	10.0	12.0	17.0	10.0	7.0	10.0	10.0	18.0
	20	9.0	7.0	9.0	10.0	16.0	9.0	7.0	9.0	11.0	16.0	9.0	6.0	9.6	8.0	16.0
	15	8.0	6.0	8.0	8.0	15.0	8.0	6.0	8.0	9.0	15.0	7.0	5.0	9.0	7.0	15.0
	10	6.0	4.7	7.0	7.0	13.0	6.0	4.0	7.0	8.0	13.0	6.0	5.0	7.0	5.2	13.0
Prevalentes (n=3189)	5	4.0	3.0	5.0	4.0	11.0	4.0	3.0	5.0	5.7	10.0	4.0	3.0	5.0	3.0	11.0
	n	3189	1450	792	329	618	1972	858	527	202	385	1214	591	265	125	233
	95	30.0	22.0	19.9	30.3	31.0	28.0	24.0	26.0	26.8	31.0	28.0	23.4	25.7	26.7	32.0
	90	28.0	21.0	17.8	27.0	30.0	25.0	22.0	23.0	25.0	29.0	25.0	21.0	21.0	24.0	30.0
	85	26.0	20.0	17.0	26.0	28.0	23.0	20.0	21.0	23.0	28.0	23.0	19.0	19.1	22.0	28.0
	80	24.0	18.6	17.0	25.0	27.0	21.0	18.0	19.0	20.4	26.0	21.0	17.0	18.0	20.8	27.0
	75	23.0	17.2	16.5	24.0	25.0	20.0	17.0	18.0	19.0	26.0	20.0	16.0	17.0	20.0	26.0
	50	18.0	13.0	12.0	17.0	21.0	15.0	12.0	14.0	15.0	21.0	14.0	11.0	14.0	15.0	23.0
	25	14.0	8.8	9.5	12.0	17.0	10.0	8.0	10.0	11.8	17.0	9.0	7.0	10.0	10.0	18.0
	20	12.0	8.0	8.0	10.6	16.0	9.0	7.0	9.0	10.0	16.0	8.0	6.0	9.2	8.2	15.0
	15	11.0	6.0	7.3	8.7	15.0	8.0	6.0	8.0	9.0	15.0	7.0	5.0	9.0	7.0	15.0
	10	8.0	4.7	5.2	5.8	13.3	6.0	4.0	7.0	7.3	13.0	6.0	5.0	7.0	6.0	13.0
Incidentes (n=526)	5	5.4	2.4	2.3	3.0	10.2	4.0	3.0	5.0	5.0	10.0	4.0	3.0	5.0	3.0	11.0
	n	526	106	41	57	322	338	65	29	31	213	183	41	12	26	104
	95	28.0	24.0	26.0	26.5	31.0	30.0	21.7	19.0	33.8	31.0	30.0	22.9	21.2	26.6	32.0
	90	25.0	21.0	22.0	24.0	29.0	28.0	20.4	17.0	29.4	30.0	28.0	21.0	21.1	25.3	29.5
	85	23.0	20.0	20.0	22.5	28.0	26.0	19.1	17.0	27.0	28.0	26.0	20.0	19.2	22.0	29.0
	80	21.0	18.0	19.0	20.0	27.0	24.0	17.0	17.0	26.6	26.0	25.0	19.0	17.8	20.2	27.0
	75	20.0	17.0	18.0	20.0	26.0	22.0	17.0	16.0	26.0	24.0	24.0	18.5	17.0	18.2	26.0
	50	14.0	12.0	14.0	15.0	22.0	18.0	12.0	13.0	18.0	20.0	18.0	15.0	11.0	13.5	21.5
	25	10.0	8.0	10.0	11.0	17.0	14.0	8.5	8.0	14.0	16.0	12.0	8.5	10.0	8.8	17.0
	20	9.0	7.0	9.0	10.0	16.0	13.0	6.2	8.0	14.0	15.8	11.8	8.0	9.6	6.2	17.0
	15	7.0	6.0	8.0	8.0	15.0	11.0	5.9	6.5	12.6	15.0	10.0	7.3	8.8	4.0	15.8
	10	6.0	4.1	7.0	7.0	13.0	8.0	3.0	5.0	9.4	14.0	8.0	5.0	6.2	3.0	12.5
	5	4.0	3.0	5.0	4.0	11.0	6.0	2.0	2.0	7.2	9.7	5.0	3.2	5.0	1.7	10.2

Notas

Tabla 7.3.- Baremos para la subescala problemas sociales de HoNOS para la muestra total, muestra incidentes y prevalentes y para hombres y mujeres según el tipo de recurso asistencial

		TOTAL					HOMBRES					MUJERES				
		Total	CSM	HD	TAC	H	Total	CSM	HD	TAC	H	Total	CSM	HD	TAC	H
Total (n=3715)	n	3715	1556	833	386	940	2310	923	556	233	598	1397	632	277	151	337
	95	14.0	11.0	11.0	13.0	15.0	14.0	11.0	11.0	13.0	15.0	14.0	11.0	12.0	13.4	15.0
	90	13.0	10.0	10.0	12.0	14.0	13.0	10.0	10.0	12.0	14.0	13.0	9.7	10.0	12.0	15.0
	85	12.0	9.0	9.0	11.0	14.0	12.0	9.0	9.0	11.0	14.0	12.0	8.0	9.0	11.0	14.0
	80	11.0	8.0	8.0	10.0	13.0	11.0	8.0	8.0	10.0	13.0	11.0	8.0	9.0	10.0	14.0
	75	10.0	7.0	8.0	9.0	13.0	10.0	8.0	8.0	9.5	13.0	10.0	7.0	8.0	9.0	13.0
	50	6.0	5.0	6.0	7.0	11.0	7.0	5.0	6.0	7.0	11.0	6.0	5.0	6.0	6.0	11.0
	25	4.0	3.0	4.0	4.0	8.5	4.0	3.0	4.0	4.5	8.0	3.0	2.0	4.0	3.0	8.5
	20	3.0	2.0	3.0	3.0	8.0	3.0	2.0	3.0	4.0	7.0	3.0	2.0	3.0	3.0	8.0
	15	3.0	2.0	3.0	3.0	7.0	3.0	2.0	3.0	3.0	7.0	2.0	1.0	3.0	2.0	6.7
	10	2.0	1.0	2.0	2.0	5.0	2.0	1.0	2.0	2.0	5.0	2.0	1.0	2.0	1.0	5.0
	5	1.0	0.0	1.0	0.0	4.0	1.0	0.0	1.0	0.0	1.0	1.0	0.0	1.9	0.0	3.0
Prevalentes (n=3189)	n	3189	1450	792	329	618	1972	858	527	202	385	1214	591	265	125	233
	95	13.0	11.0	11.0	13.0	15.0	13.0	11.0	11.0	13.0	15.0	14.0	11.0	12.0	13.7	15.0
	90	12.0	10.0	10.0	12.0	14.0	12.0	10.0	10.0	12.0	14.0	12.0	9.0	10.0	12.0	15.0
	85	11.0	9.0	9.0	11.0	14.0	11.0	9.0	9.0	10.6	14.0	11.0	8.0	9.0	11.0	14.0
	80	10.0	8.0	8.0	10.0	13.0	10.0	8.0	8.0	10.0	13.0	10.0	8.0	8.8	10.0	14.0
	75	9.0	7.0	8.0	9.0	13.0	9.0	7.0	8.0	9.0	13.0	9.0	7.0	8.0	9.0	13.0
	50	6.0	5.0	6.0	7.0	11.0	6.0	5.0	6.0	7.0	11.0	6.0	5.0	6.0	6.0	11.0
	25	4.0	3.0	4.0	4.0	8.0	4.0	3.0	4.0	4.0	8.0	3.0	2.0	4.0	4.0	8.0
	20	3.0	2.0	3.0	3.0	7.0	3.0	2.0	3.0	3.6	7.0	3.0	2.0	3.0	3.0	8.0
	15	2.0	2.0	3.0	2.5	6.0	3.0	2.0	3.0	3.0	6.0	2.0	1.0	3.0	2.0	6.0
	10	2.0	1.0	2.0	2.0	5.0	2.0	1.0	2.0	2.0	5.0	1.0	1.0	2.0	1.0	5.0
	5	1.0	0.0	1.0	0.0	4.0	1.0	0.0	1.0	0.0	4.0	1.0	0.0	2.0	0.0	3.7
Incidentes (n=526)	n	526	106	41	57	322	338	65	29	31	213	183	41	12	26	104
	95	15.0	11.0	11.9	14.0	15.0	15.0	11.0	9.5	14.4	15.3	14.0	11.0	12.0	13.6	15.0
	90	14.0	11.0	9.6	14.0	14.0	14.0	10.4	8.0	14.0	14.0	14.0	11.0	12.0	12.3	14.0
	85	13.0	9.0	8.0	13.0	14.0	14.0	9.0	8.0	14.0	14.0	13.0	10.7	12.0	11.0	14.0
	80	13.0	9.0	8.0	12.0	14.0	13.0	8.8	8.0	13.6	14.0	13.0	9.0	10.8	9.6	13.0
	75	12.0	8.0	7.5	12.0	13.0	12.0	8.0	7.0	13.0	13.0	12.0	8.5	9.5	9.0	13.0
	50	9.0	6.0	5.0	8.0	11.0	10.0	6.0	5.0	10.0	11.0	9.0	6.0	5.0	6.0	11.5
	25	6.0	3.0	3.0	5.0	9.0	6.0	3.0	3.0	6.0	9.0	6.0	4.0	3.0	2.0	9.0
	20	5.0	2.4	3.0	4.0	8.0	5.0	2.0	3.0	5.4	8.0	4.8	3.0	2.2	1.4	8.0
	15	4.0	2.0	2.0	2.7	7.0	4.0	1.9	2.0	5.0	7.0	3.0	2.0	1.0	0.0	7.8
	10	3.0	0.7	1.2	1.0	6.0	3.0	0.0	2.0	4.2	6.0	2.0	1.2	0.3	0.0	5.5
	5	1.0	0.0	0.1	0.0	3.2	2.0	0.0	0.5	2.8	3.7	1.0	0.0	0.0	0.0	2.2

Notas

En las Tablas 7.2 y 7.3 se presentan los baremos para la escala total y para la subescala de problemas sociales de HoNOS para la muestra total, muestra incidentes y prevalentes y para hombres y mujeres según el tipo de recurso asistencial.

Asimismo, en la Tabla 7.4 se presentan las PMC y los PC para dos niveles de significación distintos dependiendo del criterio de decisión a seleccionar, nivel de significación .05 y .025, para la escala total de HoNOS y para la subescala 4 de problemas sociales según sexo y tipo de recurso asistencial. La combinación de ambos criterios permite conocer si el cambio de los usuarios en cada una de las subescalas es estadísticamente significativo y clínicamente relevante.

Tabla 7.4.- Puntos de Corte y dif, mínima total HoNOS y subescala social

		Cut-off					Dif Min Cam	
		PC	A _{1d}	A _{Z=1.65}	A _{Z=1.96}	A _{p5}	DM _{Z=1.96}	DM _{Z=1.65}
Total HoNOS	Total	16.56	8.12	3.39	0.84	4.00	10.35	8.70
Rango: 0 - 48	Hombres	16.60	8.34	3.68	1.16	4.00	10.32	8.70
	Mujeres	16.44	7.75	2.93	0.33	4.00	10.10	8.50
	CSM	12.72	5.97	1.75	-0.52	3.00	9.05	7.60
	CD	13.19	8.19	4.25	2.13	5.00	9.69	8.20
	TAC	16.05	8.72	4.33	1.96	4.00	11.42	9.60
	H	21.12	14.99	10.94	8.76	11.00	11.57	9.70
Subescala IV	Total	7.68	2.87	0.28	-1.11	1.00	5.60	4.70
Problemas	Hombres	7.77	3.01	0.46	-0.92	1.00	5.68	4.80
sociales	Mujeres	7.46	2.62	0.00	-1.41	1.00	5.41	4.60
Rango: 0 - 16	CSM	5.44	1.89	-0.26	-1.41	0.00	4.84	4.10
	CD	5.58	2.84	0.87	-0.19	1.00	5.26	4.40
	TAC	7.15	3.01	0.56	-0.76	0.00	6.24	5.30
	H	10.53	7.05	4.80	3.59	4.00	5.99	5.00

Notas: PC – Punto de Corte; DifMinCam – Diferencia Mínima de Cambio; C – ; A_{1d} – 1 Desviación Estandar; A_{Z=1.65} – Alpha .025
A_{Z=1.96} – Alpha 0.05; A_{p5} – percentil 5; DM_{Z=1.96} – Diferencia mínima Alpha 0.05; DM_{Z=1.65} – Diferencia mínima Alpha .025

DISCUSIÓN

Capítulo 8

Discusión General

8.1.- Hallazgos principales

La escala Health of the Nation Outcome Scales - HoNOS - es actualmente uno de los instrumentos de referencia de los protocolos de valoración clínica en los servicios de salud mental a nivel internacional. En la revisión sistemática presentada en esta tesis, queda recogida, por un lado, la gran trayectoria clínica que tiene la escala HoNOS en los servicios de salud mental, y por otro, los numerosos estudios que presentan sus propiedades psicométricas. Si bien existen revisiones sistemáticas previas de la escala (Pirkis et al., 2005), ninguna recoge en su totalidad la bondad psicométrica publicada de los estudios ni aporta una integración de los valores de fiabilidad, validez, ni de datos clínicos presentados en dichos estudios.

Los resultados de la revisión sistemática realizada reflejan la sólida implementación de la Evaluación de Resultados (ROA) con la escala HoNOS en Reino Unido, Australia y Nueva Zelanda, siendo, además, ampliamente utilizada también en Europa (Pirkis et al 2005; Delaffon, 2012; James, 2018; Harris, 2018; Burgess, 2017; Agius, 2010; Dalsbø, 2014). La revisión recoge también su uso frecuente en España (Gisbert & Cid-Colom, 2010; Frades et al., 2011; Osta-Samanes et al., 2012; Garay-Arostegui et al., 2014; Escalada et al., 2014, 2015a, 2015b; Ribas et al., 2016) por los clínicos para la valoración de la salud global en los

distintos ámbitos asistenciales de la salud mental, a pesar de no contar con la validación española.

El objetivo principal de esta tesis ha consistido en realizar la validación de la versión española de la escala HoNOS (Uriarte, et al., 1999) a través de la Teoría Clásica de los Test (TCT) y la Teoría de Respuesta al Ítem (TRI). Para satisfacer este objetivo, el estudio se dividió en dos partes. Primero, se realizó un estudio piloto para identificar las características de adecuación psicométrica con una muestra aleatoria estratificada de 281 pacientes de la Red de Salud Mental de Bizkaia (RSMB). Y segundo, se analizó la cohorte casos y episodios registrados de la RSMB entre los años 2011 y 2020. Además de ofrecer la bondad psicométrica de la escala, esta tesis también pretende ofrecer baremos y valores de referencia de cambio clínico.

La consistencia interna de la escala valorada a través del coeficiente Alpha de Cronbach es aceptable, obteniendo un índice de .70 para la muestra del estudio piloto, y un índice ligeramente superior ($\alpha = .73$) para la cohorte de casos. Ambos valores son superiores a la fiabilidad del estudio original (.65) (Wing et al., 1998; Orrell, et al., 1999). La fiabilidad obtenida se encuentra dentro del rango reportado en la integración de datos de fiabilidad obtenidos en la revisión sistemática internacional, siendo el rango (.45 a .86), y una media de .77 (Iglesias, et al., 2019), así como dentro del rango (.59 a .76) ofrecido por la revisión de Pirkis, et al., (2005) de índices de fiabilidad de los estudios sobre la escala HoNOS.

Sin embargo, las bajas inter-correlaciones entre los ítems de las dimensiones propuestas originalmente tienen una consistencia interna relativamente pobre en el estudio piloto, especialmente la dimensión de deterioro (.10), problemas conductuales (.35) y problemas clínicos (.45), salvo la dimensión social que tiene un alfa mayor (.68). Estos valores son ligeramente más altos en la cohorte de casos (.24, .46, .37 y .73, respectivamente). Estos valores de alfa de Cronbach han sido encontrados también por otros autores (Preston, 2000; McClelland et al., 2000; Newnham et al., 2009), justificando esta baja intercorrelación entre ítems a la heterogeneidad de la escala (Newnham et al., 2009) y al bajo número de ítems que componen cada dimensión.

La fiabilidad aumentaría tan solo en el caso de eliminar el ítem 5 (.74), aunque no sería un aumento sustancial. Este resultado lo encontramos también en el estudio de Dogmanas (2014) y en el de Lovaglio y Monzani (2011). Eliminar el ítem 5 (discapacidad física) no sería aconsejable, ya que tiene sentido e importancia clínica para la escala. Además, esta modificación supondría convertir la escala HoNOS en una versión diferente a la original y no

sería así posible realizar comparaciones internacionales, ni capitalizar la importante experiencia acumulada en ella (Dogmanas, 2014; Trauer y Buckingham, 2006).

En referencia a la validez discriminante, las diferencias por centro asistencial en la gravedad de la enfermedad estarían indicando la capacidad del HoNOS para evaluar este constructo. De hecho, como cabría esperar, teniendo en cuenta el tipo de intervención llevado a cabo por cada uno de los recursos, las medias más altas con diferencia han sido recogidas por los usuarios ingresados en los hospitales psiquiátricos, mientras que las más bajas se registran en los centros ambulatorios de Salud Mental. Estos resultados los encontramos en ambas muestras estudiadas.

Este resultado es el esperado dado que los pacientes internados poseen mayores niveles de gravedad frente a otros grupos de pacientes atendidos en otros dispositivos (por ejemplo, centro de día). Pirkis et al (2005) encontraron que la escala HoNOS discrimina mejor pacientes que poseen mayor nivel de severidad. La población atendida en el Programa de Trastorno Mental Grave de Osakidetza abarca diagnósticos con distinto nivel de severidad, por lo que los resultados pueden no ser tan marcados (sin tener en cuenta el tipo de dispositivo). Pirkis et al (2005) y Dogmanas (2014) encontraron resultados similares, con puntajes más altos en pacientes con diagnósticos de consumo de drogas y pacientes psicóticos.

En lo que se refiere a la validez concurrente, la escala EEAG que mide el funcionamiento general de la persona y el EuroQol que evalúa la calidad de vida, correlacionan significativamente de forma negativa con las puntuaciones de las distintas dimensiones del HoNOS, exceptuando el EuroQol con las dimensiones de problemas conductuales y deterioro. Es decir, a un peor funcionamiento o calidad de vida (mayores puntuaciones en estas escalas) se asocia a un peor estado de salud o mayor gravedad de la enfermedad en las puntuaciones del HoNOS. Por el contrario, el HoNOS y la escala de Impresión Clínica Global correlacionan de forma positiva y estadísticamente significativa, o lo que es lo mismo a mayor mejoría menor gravedad en las distintas dimensiones.

Los resultados de los análisis factoriales confirmatorios pretenden dar respuesta a la controversia existente con relación a la estructura factorial del instrumento, y así poder esclarecer también la unidimensionalidad del mismo. Por ello, se han probado los distintos modelos, resultando ser los modelos bifactoriales los que tienen los índices más ajustados. Entre ellos el modelo bifactorial de cuatro factores correlacionados de Wing et al (1998) y el de cuatro factores correlacionados de Newnhan et al. (2009) son los que mejores índices

aportan, no habiendo diferencias claras entre ambos modelos. Por tanto, teniendo en cuenta la amplia utilización del instrumento en su estructura original, aceptar la estructura de Wing podría ser la opción más adecuada. Sin embargo, conceptualmente también podría incluirse la propuesta de Newnhan et al (2009), ya que los cuatro factores que se proponen (problemas de conducta y antisociales, deterioro físico o psicológico, síntomas y problemas socioeconómicos) han sido agrupados por clínicos y posteriormente probados mediante análisis estadísticos donde se ha encontrado una mayor consistencia interna en tres de los factores, una mejor sensibilidad al cambio y un mejor ajuste en los índices del análisis factorial confirmatorio (Newnham et al., 2009). En este sentido, ambas propuestas cumplen criterios de adecuación matemática y clínica, ya que los dos modelos han sido propuestos por expertos clínicos. No obstante, se considera que la estructura factorial original es la opción que cuenta con mayor trayectoria, ya que ha sido usada extensamente en numerosos estudios tanto clínicos (Jury et al., 2018) como psicométricos (Prabhu & Oakley-Browne, 2008).

Lo que sí permite establecer los resultados de la presente tesis es la mayor precisión del modelo bifactorial frente a los modelos simples. Esta estructura bifactorial con cinco factores indica que los ítems distribuyen sus cargas factoriales tanto por dimensiones como en un factor global. Es decir, las puntuaciones que se obtienen a través del HoNOS pueden ser interpretadas tanto por las cuatro dimensiones (problemas conductuales, deterioro, problemas clínicos y problemas sociales), como por una puntuación total de gravedad. Este modelo es acorde al procedimiento propuesto en el estudio original para interpretación de la herramienta HoNOS (Wing et al., 1998).

Los resultados de las medidas de tendencia central de los ítems indican un efecto suelo en los ítems 1 (conducta hiperactiva, agresiva, disruptiva o agitada), 2 (autolesiones no accidentales), 3 (consumo problemático de alcohol o drogas) y 11 (problemas con las condiciones de vida). Encontramos también asimetrías marcadas especialmente en los tres primeros ítems. Estos resultados los encontramos también en el estudio piloto, donde encontramos efecto suelo en estos ítems y también en el ítem 12 (problemas en relación con la ocupación y las actividades). Los pacientes que presentan valores bajos en la escala HoNOS tienen menos oportunidades de poder mostrar un cambio que aquellos pacientes que presentan puntajes iniciales más altos (Pirkis, Burgess, Kirk, et al., 2005). Las puntuaciones obtenidas en las dos muestras (incidentes-prevalentes) son bajas, por lo que no hay tamaños del efecto marcados entre ambas (prevalentes-incidentes). Debido a este efecto suelo, un bajo porcentaje de pacientes de la muestra tenía la posibilidad de presentar cambio y mejoría fiable.

Resultados similares se recogen en el estudio de Dogmanas (Dogmanas, 2014). Por otra parte, alcanzar la estabilidad de los síntomas en los casos de TMG podría considerarse un éxito terapéutico debido a la cronicidad de la enfermedad.

Una hipótesis al funcionamiento de estos ítems (1, 2, 3, 11 y 12) sería la falta de ajuste psicométrico debido a las características de la muestra. Los pacientes con Trastorno Mental Grave de la muestra tienen un largo recorrido en la Red, y a pesar de tener sintomatología crónica, logran alcanzar la estabilidad en la RSMB.

Los resultados encontrados reflejan una falta de relación entre los ítems. Esto explica la necesidad de subescalas, pero también la hiperdimensionalidad del instrumento. Los ítems 5 (problemas por enfermedad física o discapacidad) y 7 (problemas con el humor depresivo) no ajustan psicométricamente en la escala, pero tienen un sentido clínico incluirlos. El instrumento está compuesto por ítems con una significación/necesidad clínica. Es el caso también del ítem 2 (autolesiones), ya que en su origen era necesario incluir esta variable para valorar riesgo de suicidio en la población y su prevención.

8.2.- Aspectos destacados

Este estudio se realizó en la práctica rutinaria de la Red de Salud Mental de Bizkaia en funcionamiento, en lugar de con equipos de investigación seleccionados o vinculados a un programa de investigación, por lo que los resultados pueden considerarse pruebas basadas en el funcionamiento clínico y, por lo tanto, generalizables a la práctica cotidiana. Los casos analizados son, por lo tanto, representativos de la población que recibe tratamiento en los diferentes dispositivos de la red de salud mental pública de Bizkaia (Hospitales, Centros de Salud Mental, TAC y Centros de Día).

En esta tesis se recogen también los datos de baremación para la escala HoNOS. Los resultados obtenidos en este estudio se consideran especialmente útiles debido a la inexistencia de unas puntuaciones baremadas para la escala HoNOS, ya que permitirán a futuros estudios que apliquen esta medida comparar las puntuaciones en función del tipo de muestra que han recabado, y saber si los participantes de su estudio se encuentran en el área de población funcional o disfuncional. Estos baremos aportan datos para usuarios según sexo (hombres o mujeres) y también según dispositivo asistencial (Hospital, Tratamiento Asertivo Comunitario-TAC, Centro de Salud Mental-CSM y Centro de día-CD).

Además de los datos de baremación, se ofrecen también valores de referencia de cambio clínico y puntos de corte. Para la obtención de los datos que permiten la decisión del terapeuta se ha aplicado las propuestas diseñadas por Jacobson y Truax (1991), ofreciéndose puntuaciones de corte que delimitan el punto a partir del cual se puede decidir si la intensidad o frecuencia sintomática reflejada por un paciente puede ser considerada como de riesgo o no riesgo clínico. El punto de corte entre ambas poblaciones permite reducir el intervalo de puntuaciones en las que ambas se solapan, lo cual genera confusión respecto a la posición de una persona. Por otra parte, se ofrece, igualmente, la magnitud de la diferencia de puntuaciones que sería necesario experimentar tras el tratamiento para considerar si el cambio observado es por azar o atribuible a la intervención recibida.

Asimismo, permite conocer el resultado de una intervención a un individuo en particular sin necesidad de contar con una muestra amplia o tener una muestra normativa de referencia, ya que se podría comprobar el avance de cada individuo solamente con los datos que ofrece él mismo y los reflejados en este estudio. En base a estas puntuaciones de referencia, según las puntuaciones posteriores obtenidas, los pacientes se pueden clasificar como “recuperado”, “mejorado”, “no cambio” o “deteriorado”.

Suponemos el caso de una mujer que obtiene una puntuación inicial de 17 y que acude al centro de salud mental. La puntuación obtenida se sitúa en el percentil 75, lo cual indica que el 75% de las mujeres que acuden al CSM tienen una puntuación inferior y el 25% superior. Estos valores sugieren que se encuentra en una situación clínica moderada. Tras un tiempo de tratamiento, se le aplica de nuevo la escala HoNOS y obtiene una puntuación de 11, lo que le situaría en el percentil 50, indicando que ha alcanzado una puntuación inferior al 50% de las mujeres atendidas en CSM. La puntuación de 11 obtenida en la valoración post-test es inferior a la puntuación de 12.72 que es el punto de corte establecido para una desviación estándar. Sin embargo, la diferencia entre ambas puntuaciones es de 6 puntos y, considerando que la puntuación mínima de cambio calculada a través de la metodología propuesta por Jacobson y Truax (1991) es de 9.05, el clínico puede observar que este cambio no es estadísticamente significativo. Por tanto, la valoración clínica de esta mujer ha de considerarse adecuada y dentro de la funcionalidad. Sin embargo, si bien se aprecia progreso al superar el punto de corte, la magnitud de cambio no es lo suficientemente importante ($6 < PMC=9.05$). Por lo tanto, este cambio debe considerarse como probablemente debido al azar. En este caso, el profesional clínico, basándose en la combinación de ambos criterios de

decisión, puede valorar que se trata de un cambio no significativo hacia la funcionalidad en la recuperación, un cambio clínicamente adecuado, pero no estadísticamente significativo.

Otra mujer, atendida también en el CSM, ha obtenido una puntuación inicial de 21 para la escala total, lo cual la sitúa en el percentil 90. Esto supone una valoración clínica severa, ya que esta puntuación es superior a la del 90% de mujeres que acuden también a un CSM. Tras un tiempo de tratamiento en la Red, se le aplica de nuevo la escala HoNOS y obtiene una puntuación de 10, situándose en el percentil 50. La mejora alcanzada por esta paciente ha sido de 11 puntos, por lo tanto, ha superado la puntuación mínima de cambio que es de 9.05 calculada por los algoritmos propuestos por Jacobson y Truax para su grupo de referencia. Por otra también, también ha logrado una puntuación inferior al punto de corte ($PC = 12.72$) calculado para una desviación estándar. Esta mujer se halla en una situación idónea de recuperación, ya que los niveles de valoración clínica se consideran adecuados y dentro de la funcionalidad. Por un lado, supera el punto de corte y, por otro, la magnitud de cambio ha sido tal que podemos atribuirlo al tratamiento recibido ($10 > PMC=9.05$). En este caso, el profesional clínico, basándose en la combinación de ambos criterios de decisión, puede valorar que se trata de un cambio significativo hacia la funcionalidad en la recuperación; un cambio tanto clínicamente adecuado como estadísticamente significativo.

Ambas informaciones pueden considerarse de enorme interés pues ofrecen a un profesional clínico no experto en estadística una referencia para la valoración de casos individuales. Sólo le bastaría con aplicar la escala al paciente antes y después de la intervención, realizar una simple sustracción entre la medida post-test y la pre-test y comparar dicha diferencia con la magnitud que se ha calculado. Si la magnitud de cambio obtenida es mayor que la reflejada en los datos de referencia aportados, entonces el clínico podrá afirmar que sí se ha producido un cambio significativo tras su intervención. Por otro lado, si la puntuación postratamiento es inferior al punto de corte referenciado entonces el clínico podrá afirmar que además de haber producido un cambio éste ha sido en la dirección esperada de mejoría de salud y, por tanto, podrá tomar la decisión de alta terapéutica con un criterio externo además del basado en su propia experiencia. La ventaja de esta metodología es que es replicable -dos evaluadores externos llegarían a las mismas conclusiones-, y se hace eco de las aportaciones realizadas por los propios pacientes. En la revisión internacional encontramos estudios que también aplican esta metodología y ofrecen puntos de corte y valores de referencia para el cambio clínico con la escala HoNOS (Dogmanas, 2014; Maxwell et al., 2019; Parker, et al., 2020).

8.3.- Limitaciones

El estudio piloto se planificó en aras a obtener una muestra representativa de la población atendida en la Red de Salud Mental Pública de Bizkaia. Sin embargo, los datos obtenidos sólo permiten su generalización, precisamente, a la población de personas con enfermedad mental grave en tratamiento de Bizkaia. La enfermedad mental grave por definición es una entidad crónica que, entre las personas en tratamiento, tiende a estabilizarse en sus manifestaciones sintomáticas. Es por esta razón que las valoraciones del HoNOS tienden a mostrar en este estudio bajas puntuaciones, indicando dicha estabilidad dentro de su cronicidad. Es decir, la aplicación de los criterios de inclusión y la selección aleatoria de los participantes han contribuido a que la muestra del estudio piloto se caracterice por ser un grupo de enfermos estables, lo cual produce un sesgo de selección que implica que las características de comportamiento psicométrico del HoNOS sólo sean válidas para este grupo de pacientes. Al igual que en el estudio piloto, las bajas puntuaciones obtenidas en el estudio poblacional de la RSMB señalan la estabilización de las manifestaciones sintomáticas dentro de la cronicidad propia de la enfermedad mental grave. Gran parte de la muestra (85.8%) tiene un recorrido en la RSMB superior a 1 año, por lo que están estabilizados a través de los diferentes servicios/centros de la red.

Otra limitación del estudio la encontramos en la aplicación de la escala. Las bajas puntuaciones encontradas pueden deberse a la experiencia de los profesionales que aplican el protocolo de valoración, ya que encontramos que los profesionales que han realizado las evaluaciones tienden a puntuar bajo, tal vez por acomodación al síntoma tiende entonces a regular las puntuaciones. En el año 2018, James, Painter, Buckingham, et al., (2018), presentaron HoNOS-2018, donde revisaban el glosario original y proponían modificaciones para mejorar la experiencia del evaluador eliminando la ambigüedad y la inconsistencia en el glosario. Una de sus propuestas consistió en considerar las calificaciones de 0 a 1 como subclínicas, mientras que las calificaciones de 2 a 4 serían las puntuaciones que indicarían la severidad que normalmente justificaría la planificación e intervención de la atención y el tratamiento.

Por otro lado, considero una limitación la falta de estudios existentes en relación con las propiedades psicométricas de la escala HoNOS en población española. Tan solo se encuentra disponible información sobre la validez convergente respecto a la validación española de

otras escalas (ENAR-TMG - Lascorz, López, Pinedo, et al., 2016; ERP – Vargas, Quintanilla, Lozano, et al., 2013).

Por último, es importante destacar que los datos obtenidos del registro adquieren una representatividad sustantiva respecto a las personas tratadas en la Red de Salud Mental de Bizkaia, no así en otros territorios. Por tanto, los datos obtenidos serían aplicables tan solo en los servicios de Bizkaia.

8.4.- Líneas futuras

Los resultados obtenidos en la valoración de las evidencias de validez y fiabilidad psicométrica de la escala HoNOS permiten considerar esta escala como una herramienta válida y adecuada para la valoración de la gravedad clínica en personas con Trastorno Mental Grave.

En el apartado de limitaciones se señala la estabilización de los síntomas como característica de la muestra del estudio. La escala HoNOS discrimina mejor pacientes que poseen mayor nivel de severidad, por lo que sería interesante realizar un estudio en el que se valore la validez discriminante de la escala en personas en situaciones de expresión aguda de síntomas (donde se esperaría altas puntuaciones) o entre personas que estén involucradas en procesos de recuperación (por ejemplo, en programas de incorporación laboral). En este estudio podrían ofrecerse datos de referencia para cada grupo diagnóstico.

A este respecto, también sería conveniente realizar en el futuro un estudio piloto con trabajadores de diferentes recursos (Hospital, CSM, TAC y Centro de Día) y su aplicación de HoNOS con el glosario revisado (HoNOS 2018) para ver la fiabilidad en la aplicación de la escala según diferentes variables y su ajuste a los síntomas: dispositivo asistencial, experiencia de los profesionales evaluadores y severidad del usuario.

Por otra parte, sería interesante realizar un estudio donde se recoja la formación que se realiza para la aplicación de la escala y sus implicaciones psicométricas en cuanto a la fiabilidad interjueces: ¿se realiza la formación solo una vez para aprender a aplicarlo?, ¿Se repite la formación cada cierto tiempo?, ¿La formación la da siempre el mismo profesional (psiquiatras, enfermeras, psicólogos) o varía según grupo de trabajo o centro?, ¿Se ponen en consenso grupal-equipo las evaluaciones y progreso de los pacientes?. Con este estudio se podría elaborar una guía de aplicación clínica. Como referencia, está el servicio análogo de salud mental de Nueva Zelanda, Te Pou (centro nacional de salud mental, adicción y

discapacidad de Nueva Zelanda). Este centro tiene publicadas guías de aplicación para profesionales, así como formación para la aplicación de la escala HoNOS (Te Pou, 2021).

8.5.- Líneas de aplicación

Este trabajo recoge baremos, valores de referencia de cambio clínico y puntos de corte. Estos datos suponen una herramienta útil para la toma de decisiones clínicas para el profesional o equipo clínico. El desarrollo de una aplicación informática en la que se incorporen estos valores de referencia de cambio clínico permitiría automatizar el registro, seguimiento y tratamiento de cada paciente. La propia aplicación clasificaría la trayectoria de cada paciente en base a los resultados calculados, siendo estas clasificaciones como recuperado, mejorado, no cambio o deteriorado. Sería un recurso sencillo de manejar y permitiría la toma de decisiones clínicas y la evaluación de resultados de salud.

CONCLUSIONES

CAPÍTULO 9

Conclusiones

La escala Health of the Nation Outcome Scale es ampliamente utilizada a nivel internacional, incluida en los protocolos de valoración clínica en Reino Unido, Australia y Nueva Zelanda. Cuenta además con un uso extendido en los diferentes servicios de salud mental de Europa. Sus propiedades psicométricas han sido ampliamente estudiadas, pero no existía una revisión sistemática previa que recogiera todas las evidencias de adecuación psicométrica de la escala, por lo que en esta tesis se recoge una revisión sistemática internacional de las propiedades psicométricas de la escala HoNOS, realizando, a su vez, una integración meta analítica de valores como la fiabilidad o la validez convergente, así como una integración meta analítica de los datos clínicos recogidos.

La escala HoNOS se tradujo al castellano por el equipo de Uriarte et al (1999). Pese a su amplio uso clínico, no existen estudios que aporten datos de su adecuación psicométrica en población española. Por lo tanto, con este trabajo se pretende generar evidencia de la validez, fiabilidad y sensibilidad al cambio de la versión al castellano de la escala HoNOS en servicios de salud mental pública del País Vasco - Osakidetza. Estos análisis se realizaron, por una parte, con una muestra aleatoria y, por otra, con los datos de la cohorte de casos que iniciaron contacto con la Red de Salud Mental de Bizkaia (RSMB) entre los años 2011 (cuando se inició el registro de puntuaciones con el protocolo valoración inicial – Osabide Global) hasta noviembre de 2020. Por lo tanto, la evidencia psicométrica es obtenida en el contexto de la práctica clínica rutinaria de la RSMB.

El estudio de revisión internacional identifica la escala HoNOS como un instrumento de referencia para la evaluación de la gravedad clínica. Está incluida en protocolos de valoración en servicios de salud, y su uso es progresivamente mayor y más extendido. Su bondad psicométrica se encuentra extensamente estudiada y comprobada. Es una escala aceptada como válida, fiable y sensible al cambio. Se han realizado numerosos estudios donde la escala es utilizada para la validez convergente con otros instrumentos de valoración clínica; ha sido traducida a 9 idiomas y validada en 17 países. Se recoge también una fiabilidad integrada de 0.77, considerada una fiabilidad adecuada. En conjunto, los datos obtenidos en la revisión sistemática como integración del conjunto de estudios internacionales realizados sobre la escala HoNOS muestran un comportamiento adecuado para su utilización en la valoración y monitorización de procesos de enfermedad mental grave.

Por otro lado, el estudio piloto recoge las propiedades psicométricas de la escala HoNOS en una muestra representativa de la RSMB. La estructura factorial válida para este estudio es la propuesta original por los autores de la escala, y se recoge a su vez la validez convergente con los instrumentos EEAG y CGI. En cuanto a la validez interna, se obtiene un Alpha de Cronbach adecuado (.70 o .73, según la muestra utilizada en esta tesis). Estas propiedades permiten llegar a la conclusión de considerar a la escala HoNOS como un instrumento fiable y válido para la valoración gravedad clínica.

No obstante, la aplicación del modelo TRI al análisis de los ítems compromete en cierta medida su comportamiento métrico, sobre todo poniendo en cuestión la homogeneidad de la escala y los índices de discriminación.

Por otro lado, tras comprobar la adecuación psicométrica de la escala HoNOS para valorar la gravedad clínica, se proporcionan dos metodologías con el objetivo de servir como herramientas operativas para la evaluación y toma de decisiones clínicas: la baremación de la escala y el uso del Índice de Cambio Fiable. Ambas técnicas pueden ayudar al clínico a la hora de tomar las decisiones terapéuticas. Por una parte, los baremos –construidos atendiendo a la clasificación por sexo y recurso asistencial, permiten conocer y comparar la posición relativa de un usuario respecto a su grupo de referencia. Por otra parte, el Índice de Cambio Fiable (ICF) permite conocer si un usuario se encuentra por encima o por debajo del punto del corte que indica funcionalidad, además de poder conocer si el cambio que está llevando a cabo es estadísticamente significativo o no. La combinación de ambos criterios permite clasificar la evolución del individuo como: recuperado, mejorado, no cambio o deteriorado.

En conclusión, esta tesis aporta datos de la adecuación psicométrica de la versión española de la escala HoNOS, un instrumento ampliamente utilizado en los servicios de salud mental de nuestro entorno, como herramienta clínica para la valoración de la severidad global de los usuarios.

REFERENCIAS

A

- Albers, W. M. M., Roeg, D. P. K., Nijssen, Y., van Weeghel, J. and Bongers, I. M. B. (2018). Profiling of victimization, perpetration, and participation: A latent class analysis among people with severe mental illness. *PLoS ONE* 13(11) <https://doi.org/10.1371/journal.pone.0208457> (RevSist)
- Amin, S., Singh, S.P., Croudace, T., Jones, P., Medley, I., and Harrison, G. (1999). Evaluating the Health of the Nation Outcome Scales. Reliability and validity in a three-year follow-up of first- onset psychosis. *British Journal of Psychiatry*, 174, pp. 399-403. (RevSist)
- Andreas, S., Harfst, T., Dirmaier, J., Kowski, S., Koch, U., & Schulz, H. (2007). A Psychometric Evaluation of the German Version of the 'Health of the Nation Outcome Scales, HoNOS-D': On the Feasibility and Reliability of Clinician-Performed Measurements of Severity in Patients with Mental Disorders. *Psychopathology*, 40, pp. 116-125. DOI: 10.1159/000098492. (RevSist)
- Andreas, S., Harfst, T., Rabung, S., Mestel, R., Schauenburg, H., Hausberg, M., Kowski, S., Koch, U., and Schulz, H. (2009). The validity of the german version of the Health of the Nation Outcome Scales (HoNOS-D): a clinician-rating for the differential assessment of the severity of mental disorders. *International Journal of Methods in Psychiatric Research*, 19(1), pp. 50-62. (RevSist)
- Andreas, S., Harries-Hedder, K., Schwenk, W., Hausberg, M., Koch, U., and Schulz, H. (2010). Is the Health of the Nation Outcome Scales appropriate for the assessment of symptom severity in patients with substance-related disorders? *Journal of Substance Abuse Treatment*, 39, pp. 32-40. (RevSist)
- Andreas, S., Rabung, S., Mestel, R., Koch, U., Hausberg, M., and Schulz, H. (2011). Does a more specified version of the HoNOS (Health of the Nation Outcome Scales) increase psychometric properties of the inventory? *Psychopathology*, 44, pp. 261-271. (RevSist)
- Andrews, G. and Peters, L. (1994) *Measurement of consumer outcome in mental health*, in *A report to the National Mental Health Information Strategy Committee*, CRUFAD, Editor, CRUFAD: Sydney
- Andrews, G., Peters, L. and Teesson, M. (1994). *The Measurement of Consumer Outcome in Mental Health: A Report to the National Mental Health Information Strategy Committee*. Sydney: Clinical Research Unit for Anxiety Disorders.
- Anjara, S.G., Bonetto, C., Ganguli, P., Setiyawati, D., Mahendradhata, Y., Yoga, B. H., et al. (2019). Can General Practitioners manage mental disorders in primary care? A partially randomised, pragmatic, cluster trial. *PLoS ONE* 14(11), e0224724. <https://doi.org/10.1371/journal.pone.0224724> (RevSist)
- APA - Asociación Americana de Psiquiatría. (2004). *DSM-IV- Manual Diagnóstico y Estadístico de los Trastornos Mentales*. Masson.
- Apodaca, P. M. (1999). Evaluación de los resultados y del impacto. *Revista de Investigación Educativa*, 17(2), 363–377. Recuperado a partir de <https://revistas.um.es/rie/article/view/121921>
- Audin, K., Margison, F.R., Mellor-Clark, J., Barkham, M. (2001). Value of HoNOS in assessing patient change in NHS psychotherapy and psychological treatment services. *British Journal of Psychiatry*, 178, 561-566. (RevSist)
- Australian Health Ministers. (1993). *National Mental Health Plan*. Canberra: Australian Government Publishing Service.
- Australian Mental Health Outcomes and Classification Network (AMHOCN) and Community Mental Health Australia (CMHA). (2013). *National Community Managed Organisation (CMO) Outcome*

Measurement Project: Final Report to the Mental Health Information Strategy Standing Committee.
Sydney, NSW, Australia: AMHOCN.

B

- Badía, X., Bigorra, J. (2000). *La Investigación de Resultados en Salud: de la evidencia a la práctica clínica.* Xavier Badía Llach (Ed.). Barcelona: Edimac.
- Baggio, S., Kaiser, S. and Wullschlegel, A. (2024). Effect of Seclusion on Mental Health Status in Hospitalized Psychiatric Populations: A Trial Emulation using Observational Data. *Evaluation & the Health Professions*, 47(1), 3-10. doi:10.1177/01632787231164489 (RevSist)
- Barbalat, G., Van den Bergh, D. and Kossakowski, J.J. (2019). Outcome measurement in mental health services: insights from symptom networks. *BMC Psychiatry*, 19, 202. <https://doi.org/10.1186/s12888-019-2175-7> (RevSist)
- Barbato, A., Agnetti, G., D'Avanco, B., Frova, M., Guerrini, A., Tettamanti, M. (2007). Outcome of community-based rehabilitation program for people with mental illness who are considered difficult to treat. *Journal of Rehabilitation research & Development*, 44(6), pp. 775-784. (RevSist)
- Bardell-Williams, M., Eaton, S., Downey, L., Bowtell, M., Thien, K., Ratheesh, A., Killackey, E., McGorry, P. and O'Donoghue, B. (2019). Rates, determinants and outcomes associated with the use of community treatment orders in young people experiencing first episode psychosis. *International Journal of Law and Psychiatry*, 62, pp. 85-89. <https://doi.org/10.1016/j.ijlp.2018.11.007>. (RevSist)
- Barkham, M., Margison, F., Leach, C., Lucock, M., Mellor-Clark, J., Evans, C., ... McGrath, G. (2001). Service profiling and outcomes benchmarking using the CORE-OM: Towards practice-based evidence in the psychological therapies. *Journal of Consulting and Clinical Psychology*, 69, 184-196. doi:10.1037/0022-006X.69.2.184
- Barkham, M., Mellor-Clark, J., Connell, J., & Cahill, J. (2006). A core approach to practice-based evidence: A brief history of the origins and applications of the CORE-OM and CORE System. *Counselling & Psychotherapy Research*, 6, 3-15. doi:10.1080/14733140600581218
- Bartels-Velthuis, A., Ties, K., Visser, E., Arends, J., Pijnenborg, M., Wunderink, L., Jörg, F., Veling, W., Castelein, S., Kneegtering, H. and Bruggeman, R. (2020). T91. The Positive and Negative Syndrome Scale Superior to a Self-Report Questionnaire in the Pharmacotherapy Monitoring and Outcome Survey. *Schizophrenia Bulletin*, 46, Supplement 1, S266. <https://doi.org/10.1093/schbul/sbaa029.651>. (RevSist)
- Beaglehole, B., Frampton, CM., Boden, JM., Mulder, RT., and Bell, CJ. (2017). An evaluation of Health of the Nation Outcome Scales data to inform psychiatric morbidity following the Canterbury earthquakes. *Australian & New Zealand Journal of Psychiatry*, 51(11), pp. 1098-1105. DOI: 10.1177/0004867417714879 (RevSist)
- Bebbington, P., Brugha, T., Hill, T., Marsden, L., and Window, S. (1999). Validation of the Health of the Nation Outcome Scales. *British Journal of Psychiatry*, 174, pp. 389-394. (RevSist)
- Bech, P., Bille, J., Schüte, T., Søndergaard, S., Waart, S., and Wiese, M. (2003). Health of the nation Outcome Scales (HoNOS): implementability, subscale structure and responsiveness in the daily psychiatric hospital routine over the first 18 months. *Nordic Journal of Psychiatry*, 57(4), pp. 285-290. (RevSist)

- Beck, K., McCutcheon, R., Stephenson, L., Schilderman, M., Patel, N., Ramsay, R. and Howes, O. (2019). Prevalence of treatment-resistant psychoses in the community: A naturalistic study. *Journal of Psychopharmacology*, 33. 10.1177/0269881119855995. **(RevSist)**
- Benedikt, H., Lea, W., Patrik, R., and Marc, V. (2018). Coercion in substance use disorders: clinical course of compulsory admissions in a Swiss psychiatric hospital. *Swiss Medical Weekly*, 148. doi:10.4414/sm.w.2018.14644 **(RevSist)**
- Bentler, P.M. (2006). *EQS 6 Structural Equations Manual*. Encino, CA: Multivariate Software, Inc.
- Berk, M., Ng, F., Dodd, S., Callaly, T., Campbell, S., Bernardo, M., and Trauer, T. (2007). The validity of the CGI severity and improvement scales as measures of clinical effectiveness suitable for routine clinical use. *Journal of Evaluation in Clinical Practice*, 14, pp. 979-983. **(RevSist)**
- Bobes, J., Portilla, M. P., Bascarán, M. T., Sáiz, P. A., & Bousoño, M. (2002). *Banco de Instrumentos de Psiquiatría Clínica*. Barcelona: Psiquiatría Editores.
- Böckmann, V., Lay, B., Seifritz, E., Kawohl, W., Roser, P. and Habermeyer, B. (2019). Patient-Level Predictors of Psychiatric Readmission in Substance Use Disorders. *Frontiers in Psychiatry*, 10, 828. doi: 10.3389/fpsyt.2019.00828 **(RevSist)**
- Boswell, J. F., Kraus, D. R., Miller, S. D., and Lambert, M. J. (2015). Implementing routine outcome monitoring in clinical practice: Benefits, challenges, and solutions. *Psychotherapy Research*, 25(1), 6-19, DOI: 10.1080/10503307.2013.817696
- Bradley, G.M., Couchman, G.M., Perlesz, A., Nguyen, A.T., Singh, B., and Riess. (2006). Multiple-Family Group Treatment for English- and Vietnamese-Speaking Families Living with Schizophrenia. *Psychiatric Services*, 57(4), pp. 521-530. **(RevSist)**
- Branjerdporn, G., Robertson, J. and Dymond, A. (2022). Factors associated with Health of the Nation Outcomes Scales (HoNOS) in an acute young adult psychiatric unit. *International Journal of Mental Health Nursing*, 31(2), 313-325. doi:10.1111/inm.12958. **(RevSist)**
- Brann, P., Alexander, M., and Coombs, T. (2012). Routine outcome measurement in youth mental health: A comparison of two clinician rated measures, HoNOSCA and HoNOS. *Psychiatry Research*, 200, pp. 884–889. **(RevSist)**
- Brooks, R. (2000). The reliability and validity of the Health of the Nation Outcome Scales: validation in relation to patient derived measures. *Australian and New Zealand Journal of Psychiatry*, 34, pp. 504-511. **(RevSist)**
- Brower, L.S. (2003). The Ohio Mental Health Consumer Outcomes System: Reflections on a Major Policy Initiative in the US. *Clinical Psychology and Psychotherapy*, 10, 400–406. DOI: 10.1002/cpp.386
- Browne, S., Doran, M., and McGauran, S. (2000). Health of the Nation Outcome Scales (HoNOS): use in an Irish psychiatric outpatient population. *The Irish Journal of Psychological Medicine*, 17(1), 17-19. **(RevSist)**
- Bruins, J., Pijnenborg, G. H. M., Visser, E. and Castelein, S. (2021). The association of cannabis use with quality of life and psychosocial functioning in psychosis. *Schizophrenia Research*, 228, 229-234. <https://doi.org/10.1016/j.schres.2020.11.059> **(RevSist)**
- Burgess, P., Coombs, T., Clarke, A., Dickson, R., and Pirkis, J. (2012). Achievements in mental health outcome measurement in Australia: Reflections on progress made by the Australian Mental Health

Outcomes and Classification Network (AMHOCN). *International Journal of Mental Health Systems*, 6(4).

Burgess, P., Pirkis, J., and Coombs, T. (2009). Modelling candidate effectiveness indicators for mental health services. *Australian and New Zealand Journal of Psychiatry*, 43, pp. 531-538. **(RevSist)**

Burgess, P., Trauer, T., Coombs, T., McKay, R., and Pirkis, J. (2009). What does 'clinical significance' mean in the context of the Health of the Nation Outcome Scales? *Australasian Psychiatry*, 17(2), pp. 141-148. **(RevSist)**

Burgess, P.M., Harris, M.G., Coombs, T. and Pirkis, J.E. (2017). A systematic review of clinician-rated instruments to assess adults' levels of functioning in specialised public sector mental health services. *Australian & New Zealand Journal of Psychiatry*, 51(4), 338-354. doi:10.1177/0004867416688098

Buwalda, V.J.A., Draisma, S., Smit, J.H., Swinkels, J.A., Van Tilburg, W. (2011). Validering van twee meetinstrumenten voor routine outcome monitoring in de psychiatrie: de horvan-studie. *Tijdschrift Voor Psychiatrie*, 53(10), pp. 715-726. **(RevSist)**

C

Cai, W., Mueller, C., Shetty, H., et al. (2020). Predictors of cerebrovascular event reoccurrence in patients with depression: a retrospective cohort study. *BMJ Open*, 10, e031927. doi:10.1136/bmjopen-2019-031927 **(RevSist)**

Callaly, T., & Hallebone, E. L. (2001). Introducing the routine use of outcomes measurement to mental health services. *Australian health review: a publication of the Australian Hospital Association*, 24(1), 43-50. <https://doi.org/10.1071/ah010043>

Callaly, T., Trauer, T., Hantz, P. (1998). Integration of Outcome Measures into Clinical Practice. *Australasian Psychiatry*, 6(4), 188-190. doi:10.3109/10398569809084836

Callaly, T., Trauer, T., Hyland, M., Coombs, T. and Berk, M. (2011). An examination of risk factors for readmission to acute adult mental health services within 28 days of discharge in the Australian setting. *The Royal Australian and New Zealand College of Psychiatrists*, 3(19). doi: 10.3109/10398562.2011.561845 **(RevSist)**

Cella, D., Bullinger, M., Scott, C., Barofsky, I. and The Clinical Significance Consensus Meeting Group (2002). Group vs individual approaches to understanding the clinical significance of differences or changes in quality of life. *Mayo Clinic Proceedings*, 77, 384-92.

Chaplin, R., and Perkins, R. (1999). HoNOS: a cautionary tale of their use in a rehabilitation service. *Psychiatric Bulletin*, 23, pp. 20-21. **(RevSist)**

Charles, R., Bhandarkar, R. and Fulgoni, C. (2023). Using Evidence-Based Measures to Assess the Effectiveness of Residential Mental Health Rehabilitation for Adults with Dual Disabilities. *BJPsych Open*, 9, S47-S48. 10.1192/bjo.2023.183. **(RevSist)**

Chaturvedi, J., Sabbah, W., Gallagher, J. E., Turner, J., Curl, C. and Stewart, R. (2021). Hospital admissions for dental disorders in patients with severe mental illness in Southeast London: A register-based cohort study. *Eur J Oral Sci*, 129(e12752). <https://doi.org/10.1111/eos.12752> **(RevSist)**

Chicharro-Lezcano, F. (2001). Salud Mental en Euskadi. *Norte de Salud Mental*, 4(13), 131-138.

- Chieze, M., Kaiser, S., Courvoisier, D. et al. (2021). Prevalence and risk factors for seclusion and restraint in old-age psychiatry inpatient units. *BMC Psychiatry* 21(82). <https://doi.org/10.1186/s12888-021-03095-4> (RevSist)
- Chopra, P.K., Couper, J.W., Herrman, H. (2004). The assessment of patients with long-term psychotic disorders: application of the WHO Disability Assessment Schedule II. *Australian and New Zealand Journal of Psychiatry*, 38, pp. 753–759. (RevSist)
- Christensen, L. and Mendoza, J. L. (1986). A method of assessing change in a single subject: an alteration of the RC index. *Behaviour Therapy*, 17, 305–08.
- Ciarlo, J. A. (1982). Accountability revisited: the arrival of client outcome evaluation. *Evaluation and Program Planning*, 5, 31–6.
- Cocchi, A., Meneghelli, A., and Preti, A. (2008). Programma 2000: celebrating 10 years of activity of an Italian pilot programme on early intervention in psychosis. *Australian and New Zealand Journal of Psychiatry*, 42, pp. 1003-1012. (RevSist)
- Codman, E.A. (1914). The product of a hospital. *Surg Gynecol Obstet*, 18, 491–496.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2^a ed.). Hillsdale, NJ: Erlbaum.
- Colton. T (1995). *Estadística en Medicina* (10^o Edición). Barcelona: Piramide.
- Coombs, T. (2014). *What is the content and process of a comprehensive mental health nursing assessment?* University of Wollongong, New South Wales, Australia. (RevSist)
- Cranshaw, J., Seneviratne, G., Rao, R., Ogunmuyiwa, J., McMillin, R., Ntephe, C., and Dain, V. (2021). The impact of COVID-19 on an inpatient mother and baby unit: a service evaluation. *BJPsych Open*, 7(S1), pp. S316–S316. doi:10.1192/bjo.2021.833. (RevSist)
- Crook, H., Duffy, A., Nair, B., & Waters, R. (2022). Effectiveness of New Maternal Mental Health Service ‘Thrive’ in the Treatment of PTSD Symptoms Arising from Birth Trauma and Perinatal Loss. *BJPsych Open*, 8(S1), S132–S132. doi:10.1192/bjo.2022.386 (RevSist)
- Crowley, S., McDonagh, S., Carolan, D. and O’Connor, K. (2023). The Impact of a Crisis Resolution Home Treatment Team on Hospital Admission, Symptom severity and Service User Functioning over Five years. *European Psychiatry*, 19(66), Suppl 1, S467–S468. doi: 10.1192/j.eurpsy.2023.1002 (RevSist)
- Currie, S.R., Liu, P., Adamyk-Simpson, J. et al. (2020). Validation of a Comprehensive Patient Experience Survey for Addiction and Mental Health that was Co-designed with Service Users. *Community Mental Health Journal*, 56, 735–743. <https://doi.org/10.1007/s10597-019-00534-1> (RevSist)

D

- D’Avanzo, B., Battino, R.N., Gallus, S., and Barbato, A. (2004). Factors predicting discharge of patients from community residential facilities: a longitudinal study from Italy. *Australian and New Zealand Journal of Psychiatry*, 38, 619–628. (RevSist)
- D’Orta, I., Weber, K., Herrmann, F.R. et al. (2023). Determinants of clinical outcome and length of stay in acute care forensic psychiatry units. *BMC Psychiatry*, 23(264). <https://doi.org/10.1186/s12888-023-04748-2> (RevSist)
- D’Agostino, R.B., Kwan, H. (1995). Measuring effectiveness: what to expect without a randomized control group. *Medical Care*, 33, AS95-AS105.

- Dalsbø, T.K., Harboe, I., Siqveland, J., and Leiknes, K.A. (2014). *Health of the Nation Outcome Scales (HoNOS)– systematisk litteratursøk med sortert referanseliste. Notat 2014*. Oslo: Nasjonalt kunnskapssenter for helsetjenesten.
- De Beurs, E., Blankers, M., Delespaul, P., Van Duijn, E., Mulder, N., Nugter, A., Swildens, W., Tiemens, B.G., Theunissen, J., Van Voorst, A.F.A., and Van Weeghel, J. (2018). Treatment results for severe psychiatric illness: which method is best suited to denote the outcome of mental health care? *BMC Psychiatry*, 18(225). Doi: 10.1186/s12888-018-1798-4 (RevSist)
- De Beurs, E., den Hollander-Gijsman, M.E., van Rood, Y.R., van der Wee, N.J.A., Giltay, E.J., van Noorden, M.S., van der Lem, R., van Fenema, E. and Zitman, F.G. (2011), Routine outcome monitoring in the Netherlands: practical experiences with a web-based strategy for the assessment of treatment outcome in clinical practice. *Clin. Psychol. Psychother.*, 18, 1-12. <https://doi.org/10.1002/cpp.696>
- De Jong, K., and Aafjes-van Doorn, K. (2022). Routine outcome monitoring: The need for case examples. *Journal of Clinical Psychology*, 78, 1963–1972. <https://doi.org/10.1002/jclp.23441>
- De Vries, B., Pijnenborg, G. H. M., van der Stouwe, E. C. D., Visser, E., de Jong, S., Pharmacotherapy and outcome survey (PHAMOUS)-investigators, et al. (2019). “Please tell me what happened”: A descriptive study on prevalence, disclosure and characteristics of victimization in people with a psychotic disorder. *PLoS ONE* 14(7). <https://doi.org/10.1371/journal.pone.0219056> (RevSist)
- Dean, N. J., Arnaoutoglou, N. and Underwood, B. R. (2020). Effectiveness of treatment for 6813 patients with mental health conditions in Cambridgeshire: a cross-sectional study. *BJPsych Open*, 6(e30), pp. 1–7. doi: 10.1192/bjo.2020.14 (RevSist)
- Deenik, J., Tenback, D.E., Tak, E.C.P.M., Hendriksen, I.J.M., and van Harten, P.N. (2018). Improved psychosocial functioning and quality of life in inpatients with severe mental illness receiving a multidisciplinary lifestyle enhancing treatment. The MULTI study II. *Mental Health and Physical Activity*, 15, 145-152. doi: <https://doi.org/10.1016/j.mhpa.2018.10.004>. (RevSist)
- Delaffon, V., Anwar, Z., Noushad, F., Ahmed, A. S., & Brugha, T. S. (2012). Use of Health of the Nation Outcome Scales in psychiatry. *Advances in Psychiatric Treatment*, 18(3), 173–179. doi:10.1192/apt.bp.110.008029
- Department of Health. (1992). *Health of the Nation, a strategy for health in England. White Paper*. HMSO.
- Deste, G., Barlati, S., Gregorelli, M., Lisoni, J., Turrina, C., Valsecchi, P. and Vita, A. (2018). Looking through autistic features in schizophrenia using the PANSS Autism Severity Score (PAUSS). *Psychiatry Research*, 270, 764-768. <https://doi.org/10.1016/j.psychres.2018.10.074> (RevSist)
- Dhar, A., Edwards, L. and Waerea, M. (2021). A retrospective analysis comparing clinical outcome measures pre- and post- the introduction of telehealth in a community-based psychiatry clinic in a tertiary medical centre. *BJPsych Open*, 18(7), (Suppl 1), S318–S319. doi: 10.1192/bjo.2021.839 (RevSist)
- Dickens, G., Sugarman, P., y Walker, L. (2007). HoNOS-secure: a reliable outcome measure for users of secure and forensic mental health services. *Journal of Forensic Psychiatry and Psychology*, 18, 507–514.
- Dinesen Østergaard, S., Bille, J., Søltoft-Jensen, H., Lauge, N., and Bech, P. (2012). The validity of the severity–psychosis hypothesis in depression. *Journal of Affective Disorders*, 140, 48–56. (RevSist)

- Dogmanas, D. (2014). *Adaptación y validación en población uruguaya de la escala HoNOS de Resultados en Servicios de Salud Mental*. Universidad de Salamanca, Salamanca, España. (RevSist)
- Donabedian, A. (2005). Evaluating the quality of medical care. 1966. *The Milbank quarterly*, 83(4), 691–729. <https://doi.org/10.1111/j.1468-0009.2005.00397.x>
- Drakonakis, N., Stylianidis, S., Peppou, L.E. et al. (2022). Outcome of Voluntary vs Involuntary Admissions in Greece over 2 years after Discharge: A Cohort Study in the Psychiatric Hospital of Attica “Dafni”. *Community Ment Health J*, 58, 633–644. <https://doi.org/10.1007/s10597-021-00865-y> (RevSist)
- Duff, K. (2012). Evidence-Based Indicators of Neuropsychological Change in the Individual Patient: Relevant Concepts and Methods. *Archives of Clinical Neuropsychology*, 27(3), 248–261. <https://doi.org/10.1093/arclin/acr120>
- Duke, B. (2010). HoNOS in the consultation liaison psychiatry setting: is it valid? *Australasian Psychiatry*, 18(6), pp. 547–550. (RevSist)
- Duncan, B. L. (2012). The Partners for Change Outcome Management System (PCOMS): The Heart and Soul of Change Project. *Canadian Psychology / Psychologie canadienne*, 53(2), 93–104. <https://doi.org/10.1037/a0027762>

E

- Eagar, K. (2002). *An overview of health outcomes measurement in Australia. Mental Health Research & Development Strategy: Outcomes Conference*. Wellington, New Zealand
- Eagar, K., Trauer, T., and Mellsop, G. (2005). Performance of routine outcome measures in adult mental health care. *Australian and New Zealand Journal of Psychiatry*, 39, pp. 713–718. (RevSist)
- Ecob, R., Croudace, T.J., White, I.R., Evans, J.E., Harrison, G.L., Sharp, D., and Jones, P.B. (2006). Multilevel investigation of variation in HoNOS ratings by mental health professionals: naturalistic study of consecutive referrals. *International Journal of Methods in Psychiatric Research*, 13(3). (RevSist)
- Egger, S. T., Bobes, J., Theodoridou, A., Seifritz, E. and Vetter, S. (2020). Assessing the severity of psychiatric disorders using the Health of the Nation Outcome Scales: An equipercetile linking analysis. *Australian & New Zealand Journal of Psychiatry*, 54(12), 1192–1199. doi:10.1177/0004867420963725 (RevSist)
- Egger, S., Weniger, G., Prinz, S., Vetter, S., and Müller, M. (2016). Does clinical change always means the same? Comparison of different perspectives. *European Psychiatry*, 33, pp. 569. (RevSist)
- Egger, S.T., Weniger, G., Prinz, S., Vetter, S., and Müller, M. (2015). Health of the Nation Outcome Scales in a psychiatric inpatient setting: assessing clinical change. *Journal of Evaluation in Clinical Practice*, 21, pp. 236–241. (RevSist)
- Elisa, F., Davide, T., Luisa, S. M., Martina, M., Tommaso, B., Luis, S. C., ... Giulio, C. (2023). Outcome analysis on individual health budgets in mental Health: finding from the Friuli Venezia Giulia Region, Italy. *Journal of Mental Health*, 1–10. <https://doi.org/10.1080/09638237.2023.2210657>. (RevSist)
- Elvederi Murri, M., Bertelli, R., Carozza, P., et al. (2021). First-episode psychosis in the Ferrara Mental Health Department: Incidence and clinical course within the first 2 years. *Early Intervention in Psychiatry*, 15, 1738–1748. <https://doi.org/10.1111/eip.13095> (RevSist)

- Endicott, J., Spitzer, R.L., Fleiss, J.I., & Cohen, J. (1976). The Global Assessment Scale. A procedure for measuring overall severity of psychiatric disturbance. *Archives of General Psychiatry*, 33, 766-771. <https://doi.org/10.1001/archpsyc.1976.01770060086012>
- Escalada-Hernández, P., and Marín-Fernández, B. (2015). The Nursing Diagnosis Disturbed Thought Processes in Psychiatric Patients: Prevalence and Associated Characteristics. *International Journal of Nursing Knowledge*, 27(3), pp. 156-161. doi: 10.1111/2047-3095 (RevSist)
- Escalada-Hernández, P., Lazkanotegui-Matxiarena, U., Aolondriz-Urrutia, M., Alameda-Blanco, M., Flores del Redal-López, C. and Dendarrieta-Bardot, M. (2015). Diagnósticos de enfermería, intervenciones y resultados en pacientes institucionalizados con demencia. *Revista Ibero-Americana De Salud Y Envejecimiento*, 1(1), pp. 22 – 38. [https://doi.org/10.24902/r.riase.2015.1\(1\).22](https://doi.org/10.24902/r.riase.2015.1(1).22). (RevSist)
- Escalada-Hernández, P., Muñoz-Hermoso, P., González-Fraile, E., Santos, B., González-Vargas, J.A., Feria-Raposo, I., and Girón-García, J.L., and García-Manso, M. (2014). A retrospective study of nursing diagnoses, outcomes, and interventions for patients with mental disorders. *Applied Nursing Research*, 28(2), pp. 92-98. doi: 10.1016/j.apnr.2014.05.006 (RevSist)
- Esteba-Castillo, S., Torrents-Rodas, D., García-Alba, J., Ribas-Vidal, N., & Novell-Alsina, R. (2018). Translation and validation of the Spanish version of the Health of the Nation Outcome Scales for People with Learning Disabilities (HoNOS-LD). *Revista de Psiquiatría y Salud Mental*, 11(3), 141-150. <https://doi.org/10.1016/j.rpsm.2016.11.002>
- EuroQol Group (1990). EuroQol a new facility for the measurement of health-related quality of life. *Health Policy*, 16, 199-208. [https://doi.org/10.1016/0168-8510\(90\)90421-9](https://doi.org/10.1016/0168-8510(90)90421-9)
- Evans, C. E., Mellor-Clark, J., Margison, F., Barkham, M., Audin, K., Connell, J. y McGrath, G. (2000). CORE: Clinical outcome in routine evaluation. *Journal of Mental Health*, 9, 247–255

F

- Fassaert, T., Lauriks, S., van de Weerd, S., Theunissen, J., Kikkert, M., Dekker, J., Buster, M., and de Wit, M. (2014). Psychometric properties of the dutch version of the Self-Sufficiency Matrix (SSM-D). *Community Mental Health Journal*, 50, pp. 583-590. (RevSist)
- Fleminger, S., Leigh, E., Eames, P., Langrell, L., Nagraj, R., & Logsdail, S. (2005). HoNOS-ABI: a reliable outcome measure of neuropsychiatric sequelae to brain injury? *Psychiatric Bulletin*, 29(2), 53–55. doi:10.1192/pb.29.2.53
- Fonagy, P., Matthews, R., and Pilling, S. (2004) *The Mental Health Outcomes Measurement Initiative: Report from the Chair of the Outcomes Reference Group*. National Collaborating Centre for Mental Health: London.
- Fontecedro, E., Furlan, M., Tossut, D., Pascolo-Fabrizi, E., Balestrieri, M., Salvador-Carulla, L., D'Avanzo, B. and Castelpietra, G. (2020). Individual Health Budgets in Mental Health: Results of Its Implementation in the Friuli Venezia Giulia Region, Italy. *International Journal of Environmental Research and Public Health*, 17(14), 5017. <https://doi.org/10.3390/ijerph17145017> (RevSist)
- Frades, B., Ruipérez, M.A., and Moro, M. (2011). Estudio descriptivo de los pacientes atendidos en la Unidad de Media Estancia de Salud Mental del Hospital Pare Jofré durante los años 2005-2008. *Rehabilitación Psicosocial*, 8(1/2), pp. 23-31. (RevSist)

G

- Gale, T. M. and Boland, B. (2021). COVID-19 deaths in a secondary mental health service. *Comprehensive Psychiatry*, 152277(111). <https://doi.org/10.1016/j.comppsy.2021.152277>. (RevSist)
- Gale, T.M. and Boland, B. (2019). A model for predicting missing items on the Health of the Nation Outcome Scale (HoNOS). *Comprehensive Psychiatry*, 93, pp. 61–64. <https://doi.org/10.1016/j.comppsy.2019.06.005> (RevSist)
- Gallagher, J., and Teesson, M. (2000). Measuring disability, need and outcome in Australian community mental health services. *Australian and New Zealand Journal of Psychiatry*, 34, pp. 850–855. (RevSist)
- Garay-Arostegui, M., Pousa-Rodríguez, V., & Pérez-Cabeza, L. (2014). La relación entre la percepción subjetiva del funcionamiento cognitivo y el autoestigma con la experiencia de recuperación de las personas con enfermedad mental grave. *Revista de la Asociación Española de Neuropsiquiatría*, 34(123), 459-475. <https://doi.org/10.4321/S0211-57352014000300002> (RevSist)
- Gaupp, R., Walter, M., Bader, K., Benoy, C. and Lang, U. E. (2020). A Two-Day Acceptance and Commitment Therapy (ACT) Workshop Increases Presence and Work Functioning in Healthcare Workers. *Frontiers in Psychiatry*, 11, 861. doi: 10.3389/fpsy.2020.00861 (RevSist)
- Georgaca, E., Machaira, S., Stamovlasis, D., Peppou, L. E., Papachristou, C., Arvaniti, A., Samakouri, M., Stylianidis, S., Bozikas, V. P., Diakogiannis, I., Fokas, K., Garyfallos, G., Gkolia, I., Karpouza, V., Nimatoudis, I., Patsinakidis, G., Sevrin, D. and Vlachaki, A. (2023). Clinical determinants of involuntary psychiatric hospitalization: A clinical profile approach. *J Clin Psychol*, 79(9), 2081-2100. doi: 10.1002/jclp.23528 (RevSist)
- Gilbert, M.M., Chamberlain, J.A., White, C.R., Mayers, P.W., Pawsey, B., Liew, D., Musgrave, M., Crawford, K., and Castle, D.J. (2012). Controlled clinical trial of a self-management program for people with mental illness in an adult mental health service – the Optimal Health Program (OHP). *Australian Health Review*, 36, pp. 1–7. <http://dx.doi.org/10.1071/AH11008> (RevSist)
- Gilbody, S. M., House, A. O., y Sheldon, T. A. (2002a). Outcome research in mental health. *British Journal of Psychiatry*, 181, 8–16.
- Gilbody, S. M., House, A. O. and Sheldon, T. A. (2003). Outcomes measurement in psychiatry: a critical review of outcomes measurement in psychiatric research and practice. *Research Report. CRD Report (24)*. NHS Centre for Reviews and Dissemination, York, UK.
- Gisbert, C., & Cid-Colom, J. (2010). *Evaluación en unidades de rehabilitación hospitalaria: Un balance entre síntomas, funcionalismo, necesidades, expectativas y soporte comunitario*. En FEARP (Ed.), *Evaluación en rehabilitación psicosocial*, (pp. 213-231). Federación Española de Asociaciones de Rehabilitación Psicosocial (FEARP). (RevSist)
- Glover, G. (1997). The development of a new minimum data set for specialist mental health. *Health Trends*, 29, 48–51.
- Gobierno Vasco. (2023) *Estrategia de Salud Mental de Euskadi 2023-2028*. https://bideoak2.euskadi.eus/2023/11/21/news_89832/Estrategia_de_Salud_Mental_es.pdf
- Golay, P., Basterrechea, L., Conus, P., and Bonsack, C. (2016). Internal and predictive validity of the French Health of the Nation Outcome Scales: need for future directions. *Plos One*, 11(8). (RevSist)

- Goldney, RD., Fisher, LJ., and Walmsley, SH. (1998). The Health of the Nation Outcome Scales in psychiatric hospitalisation: a multicenter study examining outcome and prediction of length of stay. *Australian and New Zealand Journal of Psychiatry*, 32, pp. 199-205. **(RevSist)**
- Gonda, T., Deane, F.P., and Murugesan, G. (2012). Predicting clinically significant change in an inpatient program for people with severe mental illness. *Australian & New Zealand Journal of Psychiatry*, 46(7), pp. 651– 658. Doi: 10.1177/0004867412445527 **(RevSist)**
- Gore-Jones, V., and Dark, F. (2018). Key performance indicators and administrative data in an early psychosis service. *Early Intervention in Psychiatry*, pp. 1–6. DOI: 10.1111/eip.12759. **(RevSist)**
- Green, HJ., Frank, RM., Butwell, A., and Beck, OJ. (2007). Implementation and evaluation of brief cognitive –behavioural therapy in a mental health acute assessment and treatment service. *Behavior Change*, 24(2), pp. 87-98. **(RevSist)**
- Guy, W. (1976). *ECDEU Assessment Manual for Psychopharmacology*. Rockville: Rockville National Institute of Mental Health.
- Guy, W. (2004). *Escala de Impresión Clínica Global*. En J. Bobes, M. P. García-Portilla, M.T., Bascarán, P. A. Sáiz & M. Bousoño (Eds). Banco de Instrumentos Básicos para la Práctica de la Psiquiatría Clínica, 101-102. Barcelona: Psiquiatría Editores.

H

- Hall, J. (1996). The challenge of health outcomes. *Journal of Quality Clinical Practice*, 16, 5-15.
- Hansen, JP., Østergaard, B., Nordentoft, M., and Hounsgaard, L. (2012). Cognitive adaptation training combined with assertive community treatment: a randomized longitudinal trial. *Schizophrenia Research*, 135, pp. 105-111. **(RevSist)**
- Hansen, L., and Kingdon, D. (2006). Rating suicidality in schizophrenia: items on global scales (HoNOS and CPRS) correlate with a validated suicidality rating scale (InterSePT). *Archives of Suicide Research*, 10, 249-252. **(RevSist)**
- Happell, B., Sundram, S., Wortans, J., Johnstone, H., Ryan, R., and Lakshmana, R. (2009). Assessing Nurse-Initiated Care in a Mental Health Crisis Assessment and Treatment Team in Australia. *Psychiatric Services*, 60(11), pp. 1527-1531. **(RevSist)**
- Hariman, K., Cheng, K. M., Lam, J., Leung, S. K., Lui, S. S. Y. (2020). Clinical risk model to predict 28-day unplanned readmission via the accident and emergency department after discharge from acute psychiatric units for patients with psychotic spectrum disorders. *BJPsych Open*, 6(1), e13. doi: 10.1192/bjo.2019.97. **(RevSist)**
- Harris, M.G., Sparti, C., Scheurer, R., et al. (2018). Measurement properties of the Health of the Nation Outcome Scales (HoNOS) family of measures: protocol for a systematic review. *BMJ Open*, 8, e021177. doi:10.1136/bmjopen-2017-021177
- Health and Social Care Information Centre (HSCIS). (2014). *Mental Health Bulletin. Annual Report from MHMDS Returns 2013–14*. Retrieved 14 enero 2020 from <https://files.digital.nhs.uk/publicationimport/pub15xxx/pub15990/mhb-1314-ann-rep.pdf>
- Herdman, M., Badia, X., & Berra, S. (2001). EuroQol-5D: Una alternativa sencilla para la medición de la calidad de vida relacionada con la salud en atención primaria. *Atención Primaria*, 28(6), 425-430. [https://doi.org/10.1016/S0212-6567\(01\)70406-4](https://doi.org/10.1016/S0212-6567(01)70406-4)

- Hodgson, R.E. (2019). Evaluating the cost and clinical effectiveness of long-acting, injectable aripiprazole and paliperidone palmitate once a month in a real-world setting. *Clinico Economics and Outcomes Research*, 11, pp. 517–524. <https://doi.org/10.2147/CEOR.S191198> (RevSist)
- Hofmann, A.B., Schmid, H. M., Hofmann, L. A., Noboa, V., Seifritz, E., Vetter, S. and Egger, S. T. (2022). Impact of compulsory admission on treatment and outcome: A propensity score matched analysis. *European Psychiatry*, 65(1), e6, 1–14. <https://doi.org/10.1192/j.eurpsy.2022.4> (RevSist)
- Hong, J. S. W., Sheriff, R., Smith, K., et al. (2021). Impact of COVID-19 on telepsychiatry at the service and individual patient level across two UK NHS mental health Trusts. *Evidence Based Mental Health*, 24, pp. 161–166. doi:10.1136/ebmental-2021-300287 (RevSist)
- Hooke, G.R., and Page, A.C. (2002). Predicting Outcomes of Group Cognitive Behavior Therapy for Patients with Affective and Neurotic Disorders. *Behavior Modification*, 26(5), pp. 648-658. DOI: 10.1177/014544502236655 (RevSist)
- Hu, L. and Bentler, P.M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Struct Equ Modeling*, 6(1), 1–55. doi: 10.1080/10705519909540118
- Hugo, M. (2000). Comparative efficiency ratings between public and private acute inpatient facilities. *Australian and New Zealand Journal of Psychiatry*, 34, pp. 651–657. (RevSist)
- Hugo, M., Smout, M., and Bannister, J. (2002). A comparison in hospitalization rates between a community-based mobile emergency service and a hospital-based emergency service. *Australian and New Zealand Journal of Psychiatry*, 36, pp. 504–508. (RevSist)
- Hunter, R., Mclean, J., Peck, D., Pullen, I., Greenfield, A., Mcarthur, W., and Quinn, A., Eaglesham, J., Hagen, S., and Norrie, J. (2004). The Scottish 700 Outcomes Study: A Comparative Evaluation of the Health of the Nation Outcome Scale (HoNOS), the Avon Mental Health Measure (AVON), and an Idiographic Scale (OPUS) in adult mental health. *Journal of Mental Health*, 13(1), pp. 93-105. (RevSist)

I

- Iglesias, N., Penas, P., Macía, M., & Iraurgi, I. (2019). *Revisión sistemática internacional de las características psicométricas de la escala HoNOS (Health of the Nation Outcome Scale)*. Comunicación oral. IV Congreso Nacional de Psicología y I International Symposium on Psychological Prevention. Vitoria, España.
- Iraurgi, I. (2009a). Evaluación de resultados clínicos (I): Entre la significación estadística y la relevancia clínica. *Norte de Salud Mental*, 33, 94-108.
- Iraurgi, I. (2009b). Evaluación de resultados clínicos (II): Las medidas de la significación clínica o los tamaños del efecto. *Norte de salud mental*, 34(8), 94-110.
- Iraurgi, I. (2010). Evaluación de resultados clínicos (y III): Índices de Cambio Fiable (ICF) en la estimación del cambio clínicamente significativo. *Norte de Salud Mental*, 36(8), 105-122. [http://www.ome-aen.org/NORTE/36/11.%20Formación%20continuada%20\(1\)%20Norte36.pdf](http://www.ome-aen.org/NORTE/36/11.%20Formación%20continuada%20(1)%20Norte36.pdf)
- Irvine, A., Rawlinson, C., Bor, W., & Hoehn, E. (2021). Evaluation of a collaborative group intervention for mothers with moderate to severe perinatal mental illness and their infants in Australia. *Infant Mental Health Journal*, 42, pp. 560–572. <https://doi.org/10.1002/imhj.21922> (RevSist)

- Issakidis, C., and Teesson, M. (1999). Measurement of need for care: a trial of the Camberwell Assessment of Need and the Health of the Nation Outcome Scales. *Australian and New Zealand Journal of Psychiatry*, 33, pp. 754–759. (RevSist)

J

- Jacobs, R. (2007). *Investigating Patient Outcome Measures in Mental Health: research report for the OHE Commission on NHS Productivity*. York: Centre for Health Economics, University of York.
- Jacobs, R. (2009). *Investigating Patient Outcome Measures in Mental Health*. York: Centre for Health Economics, University of York.
- Jacobson, N. S., Follette, W. C. and Revensdorf, D. (1984). Psychotherapy outcome research: Methods for reporting variability and evaluating clinical significance. *Behaviour Therapy*, 15, 336–52.
- Jacobson, N.S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, 59, 12-19. <https://doi.org/10.1037/0022-006x.59.1.12>
- Jacobson, N.S., y Truax, P. (1991). Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, 59, 12-19.
- James, M., Painter, J., Buckingham, B., and Stewart, MW. (2018). A review and update of the Health of the Nation Outcome Scales (HoNOS). *BJPsych Bulletin*, 42, pp. 63–68. doi:10.1192/bjb.2017.17
- Jansen, L., Hunnik, F., Busschbach, J.J.V., Lijmer, J.G. (2019). Measuring outcomes on a Medical Psychiatric Unit: HoNOS, CANSAS and costs. *Psychiatry Research*, 280, 112526. <https://doi.org/10.1016/j.psychres.2019.112526>. (RevSist)
- Jefford, M., Stockler, M. R., & Tattersall, M. H. (2003). Outcomes research: what is it and why does it matter? *Internal medicine journal*, 33(3), 110–118. <https://doi.org/10.1046/j.1445-5994.2003.00302.x>
- Jochems, E.C., Duivenvoorden, H.J., van Dam, A. et al. (2018). Testing the Integral Model of treatment motivation in outpatients with severe mental illness. *Motivation and Emotion*, 42, 816–830. <https://doi.org/10.1007/s11031-018-9708-0> (RevSist)
- Jones, C.I., Bremner, S., Hooper, R. et al. (2021). Statistical analysis plan for the Early Youth Engagement in first episode psychosis (EYE-2) study: a pragmatic cluster randomised controlled trial of implementation, effectiveness and cost-effectiveness of a team-based motivational engagement intervention to improve engagement. *Trials*, 732(22). <https://doi.org/10.1186/s13063-021-05670-2> (RevSist)
- Jury, A., Lai, J., Tuason, C., Koning, A., Smith, M., Boyd, L., Swanson, C., Fergusson, D., and Gruar, A. (2018). People who experience seclusion in adult mental health inpatient services: An examination of health of the nation outcome scales scores. *International Journal of Mental Health Nursing*, 28(1), pp. 199-208. doi: 10.1111/inm.12521 (RevSist)

K

- Kazdin, A. E. (2008). Evidence-based treatment and practice: New opportunities to bridge clinical research and practice, enhance the knowledge base, and improve patient care. *American Psychologist*, 63, 146–159. doi:10.1037/0003-066X.63.3.146

- Kelly, C., Sharkey, V., Morrison, G., Allardyce, J., and McCready, R.G. (2000). Cognitive function in a catchment-area-based population of patients with schizophrenia. *British Journal of Psychiatry*, 177, 348-353. (RevSist)
- Kelly, P.J. and Deane, F.P. (2009). Relationship between therapeutic homework and clinical outcomes for individuals with severe mental illness. *Australian and New Zealand Journal of Psychiatry*, 43, pp. 968-975. (RevSist)
- Kennedy, C., and Yellowlees, P. (2002). The effectiveness of telepsychiatry measured using the Health of the Nation Outcome Scale and the Mental Health Inventory. *Journal of Telemedicine and Telecare*, 9, pp. 12-16. (RevSist)
- Kipling, R. (1901). *Just so Stories*. MacMillan.
- Kisely, S., Campbell, L.A., Cartwright, J., Cox, M., and Campbell, J. (2010). Do the Health of the Nation Outcome Scales Measure Outcome? *The Canadian Journal of Psychiatry*, 7(55). (RevSist)
- Kisely, S., Campbell, L.A., Crossman, D., Gleich, S., and Campbell, J. (2007). Are the Health of the Nation Outcome Scales a Valid and Practical Instrument to Measure Outcomes in North America? A Three-Site Evaluation Across Nova Scotia. *Community Mental Health Journal*, (43)2. DOI: 10.1007/s10597-006-9067-2. (RevSist)
- Kisely, S., Preston, N., and Rooney, M. (2000). Pathways and outcomes of psychiatric care: does it depend on who you are, or what you've got? *Australian and New Zealand Journal of Psychiatry*, 34, pp. 1009–1014. (RevSist)
- Kisely, S., Xiao, J., Crowe, E., Paydar, A., Jian, L. (2014). The effect of community treatment orders on outcome as assessed by the Health of the Nation Outcome Scales. *Psychiatry Research*, 215, pp. 574–578. (RevSist)
- Kocha, N., Wasserfallenb, J.B., Camusa, D., Gebela, S., Bonsacka, C. (2013). Implementing validated routine outcome assessment in psychiatry: a field study. *Swiss Archives of Neurology and Psychiatry*. 164(4), 134–44. (RevSist)
- Kortrijk, H. E., Mulder, C. L., van der Gaag, M. and Wiersma, D. (2012). Symptomatic and functional remission and its associations with quality of life in patients with psychotic disorder in Assertive Community Treatment teams. *Comprehensive Psychiatry*, 53(8), 1174-1180. <https://doi.org/10.1016/j.comppsy.2012.05.001>. (RevSist)
- Kortrijk, H., Schaefer, B., van Weeghel, J., Mulder, C.L., and Kamperman, A. (2019). Trajectories of patients with severe mental illness in two-year contact with Flexible Assertive Community Treatment teams using Routine Outcome Monitoring data: An observational study. *PLoS ONE* 14(1). Doi: 10.1371/journal.pone.0207680 (RevSist)
- Kortrijk, H.E., Mulder, C.L., Drukker, M., Wiersma, D., and Duivenvoorden, H.J. (2012). Duration of assertive community treatment and the interpretation of routine outcome data. *Australian & New Zealand Journal of Psychiatry*, 46(3), pp. 240–248. DOI: 10.1177/0004867411433214 (RevSist)
- Kortrijk, H.E., Mulder, N.L., Kamperman, A.M. et al. (2019). Employment Rates in Flexible Assertive Community Treatment Teams in The Netherlands: An Observational Study. *Community Mental Health Journal*, 55, 350–359. <https://doi.org/10.1007/s10597-018-0233-0> (RevSist)
- Kugathasan, P., Wu, H., Gaughran, F., Nielsen, R. E., Pritchard, M., Dobson, R., Stewart, R. and Stubbs, B. (2020). Association of physical health multimorbidity with mortality in people with schizophrenia spectrum disorders: Using a novel semantic search system that captures physical diseases in electronic

patient records. *Schizophrenia Research*, 216, pp.408-415.
<https://doi.org/10.1016/j.schres.2019.10.061> (RevSist)

Kwakernaak, S., Swildens, W.E., van Wel, T.F. et al. (2020). Symptomatic and Functional Remission in Young Adults with a Psychotic Disorder in a Rehabilitation Focused Team. *Community Mental Health Journal*, 56, 549–558. <https://doi.org/10.1007/s10597-019-00512-7> (RevSist)

L

La Tona, A., Tagini, S., Brugnera, A., Poletti, B., Aiello, E. N., Lo Coco, G., Del Piccolo, L., & Compare, A. (2023). Italian validation of the Clinical Outcomes in Routine Evaluation-10 (CORE-10): a short measure for routine outcome monitoring in clinical practice. *Research in psychotherapy (Milano)*, 26(1), 671. <https://doi.org/10.4081/ripppo.2023.671>

Lai, J., Jury, A., Long, J., Fergusson, D., Smith, M., Baxendine, S., and Gruar, A. (2019). Variation in seclusion rates across New Zealand's specialist mental health services: Are sociodemographic and clinical factors influencing this? *International Journal of Mental Health Nursing*, 28(1), pp. 288-296. (RevSist)

Lally, J., Watkins, R., Nash, S., Shetty, H., Gardner-Sood, P., Smith, S., Murray, RM., and Gaughran, F. (2018). The Representativeness of Participants with Severe Mental Illness in a Psychosocial Clinical Trial. *Frontiers in Psychology*, 654(9). (RevSist)

Lambert, M. J., Burlingame, G. M., Umphress, V., Hansen, N. B., Vermeersch, D. A., Clouse, G. C., & Yanchar, S. C. (1996). The reliability and validity of the outcome questionnaire. *Clinical Psychology & Psychotherapy*, 3, 249-258. doi:10.1002/(SICI)1099- 0879(199612)3:43.0.CO;2-S

Lambert, M. J., Hansen, N. B., & Finch, A. E. (2001). Patient-focused research: Using patient outcome data to enhance treatment effects. *Journal of Consulting and Clinical Psychology*, 69, 159 –172.

Lambert, M. J., Whipple, J. L., & Kleinstäuber, M. (2018). Collecting and delivering progress feedback: A meta-analysis of routine outcome monitoring. *Psychotherapy*, 55(4), 520–537. <https://doi.org/10.1037/pst0000167>

Lascorz, D., Lopez, V., Pinedo, C., Trujols, J., Vegué, J., and Perez, V. (2018). Estudio psicométrico de la Escala de valoración de los Niveles de Atención Requerida para personas con Trastorno Mental Grave (ENAR-TMG). *Revista de Psiquiatría y Salud Mental*, 11(3), 156-168. (RevSist)

Laugharne, R., Eaves, S., Mascas, A., Psatha, K., Dinnis, G., Trower, J., and Shankar, R. (2018). Desperately seeking outcomes: quantifying the effectiveness of community mental healthcare using Health of the Nation Outcome Scales. *The British Journal of Psychiatry*, 4, pp. 91-94. doi: 10.1192/bjo.2018.3 (RevSist)

Lauzon, S., Corbière, M., Bonin, JP., Bonsack, C., Lesage, AD., and Ricard, N. (2001). Validation de la version française du Health of the Nation Outcome Scales (HoNOS-F). *The Canadian Journal of Psychiatry*, 46, pp. 841-846. (RevSist)

Leginski, W. A., Croze, C., Driggers, J., et al. (1989). *Data Standards for Mental Health Decision Support Systems. A Report of the Task Force to Revise the Data Content and System Guidelines of the Mental Health Statistics Improvement Program*. Washington DC.

Lei, H., Barnicot, K., Maynard, E., Etherington, A., Zalewska, K., Quirk, A., Sanatinia, R., Cooper, S. J., Crawford, M. J. (2019). Factors influencing use of community treatment orders and quality of care

- that people receive: results of a national survey in England and Wales. *BJPsych Bulletin*, 11, 227-235. doi: 10.1192/bjb.2019.23 (RevSist)
- Leijala, J., Kampman, O., Suvisaari, J. et al. (2021). Daily functioning and symptom factors contributing to attitudes toward antipsychotic treatment and treatment adherence in outpatients with schizophrenia spectrum disorders. *BMC Psychiatry*, 21(37). <https://doi.org/10.1186/s12888-021-03037-0> (RevSist)
- Liemburg, E., Nienhuis, F. and Veling, W. (2020). M95. DSM-5 Clinician-Rated Dimensions of Psychosis Symptom Severity: Psychometric Properties. *Schizophrenia Bulletin*, 46, Suppl 1, S170–S171. doi: 10.1093/schbul/sbaa030.407 (RevSist)
- Listunova, L., Bartolovic, M., Kienzle, J., Jaehn, A., Grützner, T. M., Wolf, R. C., Weisbrod, M. and Roesch-Ely, D. (2020). Predictors of cognitive remediation therapy improvement in (partially) remitted unipolar depression. *Journal of Affective Disorders*, 264, 40-49. <https://doi.org/10.1016/j.jad.2019.12.006> (RevSist)
- López-Morinigo, JD., Fernandes, AC., Chang, CK., Hayes, RD., Broadbent, M., Stewart, R., David, AS., and Dutta, R. (2014). Suicide completion in secondary mental healthcare: a comparison study between schizophrenia spectrum disorders and all other diagnoses. *BMC Psychiatry* 14, pp. 213-223. (RevSist)
- Lora, A., Bai, G., Bianchi, S., Bolongaro, G., Civalenti, G., Erlicher, A., Morosini, P. (2001). The Italian version of HoNOS (Health of the Nation Outcome Scales), a scale for evaluating the outcomes and the severity in mental health services. *Epidemiologia e Psichiatria Sociale*, 10(3), 198–204. doi:10.1017/S1121189X00005339
- Lorant, V., Grard, A., Van Audenhove, C. et al. (2019). Effectiveness of Health and Social Service Networks for Severely Mentally Ill Patients' Outcomes: A Case-Control Study. *Administration and Policy in Mental Health and Mental Health Services Research*, 46, 288–297. <https://doi.org/10.1007/s10488-018-0910-x> (RevSist)
- Lorenzo-Seva, U., & Ferrando, P. J. (2013) *Factor Program for windows*. Available in: <http://psico.fcep.urv.es/utilitats/factor/index.html>
- Lorenzo-Seva, U., & Ferrando, P.J. (2019). *Factor 10. Manual of the Program*. Universidad Rovira I Virgili. Recuperado de <http://psico.fcep.urv.es/utilitats/factor>
- Lovaglio, P.G., & Monzani, E. (2012). Health of the nation outcomes scales evaluation in a community setting population. *Quality of Life Research*, 21, 1643-1653. <https://doi.org/10.1007/s11136-011-0071-9>
- Lovaglio, PG., and Monzani, E. (2011). Validation aspects of the health of the nation outcome scales. *International Journal of Mental Health Systems*, 20(5). (RevSist)
- Lovaglio, PG., and Monzani, E. (2012). Health of the nation outcome scales evaluation in a community setting population. *Quality of Life Research*, 21, pp. 1643-1653. (RevSist)
- Lund-Heimark, H., Kjølby, E., Mehlum, L., Gjestead, R., Selbæk, G., Kroken, R. A., Johnsen, E., Oedegaard, K. J. and Mellesdal, L. S. (2020). Elderly patients with no previous psychiatric history: suicidality and other factors relating to psychiatric acute admissions. *BJPsych Open*, 18, 6(4). doi: 10.1192/bjo.2020.45 (RevSist)
- Luo, W., Harvey, R., Tran, T., et al. (2016). Consistency of the Health of the Nation Outcome Scales (HoNOS) at inpatient-to-community transition. *BMJ Open*, 6, e010732. doi:10.1136/bmjopen-2015-010732 (RevSist)

Lutchman, R., Thompson, A., Tait, H., et al. (2007) In search of a standardised, comprehensive assessment of functioning. *New Zealand Journal of Occupational Therapy*, 54, 33–38.

M

Macdonald, A., Adamis, D., Craig, T., and Murray, R. (2018). Continuity of care and clinical outcomes in the community for people with severe mental illness *The British Journal of Psychiatry*. 1-6. doi:10.1192/bjp.2018.261 (RevSist)

Machova', K., Juri'čková', V., Kasparova', A., Petrova', K., Sla'dkova', B. and Svobodova', I. (2023). An evaluation of the effect of equine-facilitated psychotherapy on patients with substance use disorders. *PLoS ONE*, 18(6), e0286867. <https://doi.org/10.1371/journal.pone.0286867> (RevSist)

Maddison, P., Golay, P., and Muncer, SJ. (2018). Latent class analysis of the health of the nation outcome scales: A comparison of Swiss and English profiles and exploration of their predictive utility. *European Psychiatry*, 48, pp. 1-5. (RevSist)

Mansour, H., Mueller, C., Davis, K. A. S., Burton, A., Shetty, H., et al. (2020). Severe mental illness diagnosis in English general hospitals 2006-2017: A registry linkage study. *PLOS Medicine* 17(9), e1003306. <https://doi.org/10.1371/journal.pmed.1003306> (RevSist)

Marks, I. (1998). Overcoming obstacles to routine outcome measurement. *British Journal of Psychiatry*, 173, 281–286.

Martin-Iñigo, L., Ortiz, S., Urbano, D. et al. (2022). Assessment of the efficacy of a Crisis Intervention Team (CIT): experience in the Esplugues Mental Health Center (Barcelona). *Soc Psychiatry Psychiatr Epidemiol*, 57, pp. 2109–2117. <https://doi.org/10.1007/s00127-022-02250-w> (RevSist)

Martín-Pérez, C., Herrero-Martin, J. J., Pérez-López, G., Pedrosa-García, R., Ibañez-Casas, I., Torres-González, F. and Moreno-Küstner, B. (2019). Severe Mental Illness in Community Mental Health Care in Spain Prevalence and Related Factors. *The Journal of Nervous and Mental Disease*, 2(207). DOI:10.1097/NMD.0000000000000929 (RevSist)

Mascio, A., Stewart, R., Botelle, R., Williams, M., Mirza, L., Patel, R., Pollak, T., Dobson, R. and Roberts, A. (2021). Cognitive Impairments in Schizophrenia: A Study in a Large Clinical Sample Using Natural Language Processing. *Front. Digit. Health*, 3(711941). doi: 10.3389/fdgth.2021.711941 (RevSist)

Mavreas, V. and Fotiadou, M. (2010). *Greek version of HoNOS65*. Available at: http://www.rcpsych.ac.uk/pdf/CETC_HoNOS%2065p_tabulated_Greek%20v2

Maxwell, A., Tsoutsoulis, K., Tarur Padinjareveettil, A. M., Zivkovic, F. and Rogers, J. M. (2018). Longitudinal analysis of statistical and clinically significant psychosocial change following mental health rehabilitation. *Disability and Rehabilitation*, 41:24, pp. 2927-2939. DOI: 10.1080/09638288.2018.1482505 (RevSist)

McClelland, R., Trimble, P., Fox, ML., Stevenson MR., and Bell, B. (2000). Validation of an outcome scale for use in adult psychiatric practice. *Quality in Health Care*, 9, pp. 98-105. <https://doi.org/10.1136/ghc.9.2.98> (RevSist)

McDowell, I. (2006). *Measuring Health: A Guide to Rating Scales and Questionnaires*. Third Edition. Oxford University Press.

- McKay, R., and Coombs, T. (2015). An exploration of the ability of routine outcome measurement to represent clinically meaningful information regarding individual consumers. *Australasian Psychiatry*, 20(5), pp. 433–437. DOI: 10.1177/1039856212458813 (RevSist)
- McNaught, M., Caputi, P., Oades, LG., and Deane, FP. (2007). Testing the validity of the Recovery Assessment Scale using an Australian sample. *Australian and New Zealand Journal of Psychiatry*, 41, pp. 450–457. (RevSist)
- McPheeters, H. L. (1980). Measurement of mental health program outcomes. *New Directions for Program Evaluation*, 6, 53–63.
- Mellsop, G., Dutu, G., and El-Badri, S. (2007). CAOS contribution to understanding cultural/ ethnic differences in the prevalence of bipolar affective disorder in New Zealand. *Australian and New Zealand Journal of Psychiatry*, 41, pp. 392–396. (RevSist)
- Meroni, G., Sentissi, O., Kaiser, S. and Wullschleger, A. (2023). Treatment without consent in adult psychiatry inpatient units: a retrospective study on predictive factors. *Frontiers in Psychiatry*, 14:1224328. doi: 10.3389/fpsyt.2023.1224328 (RevSist)
- Miller, K., Walterfang, M., Randhawa, S., Scholes, A., Mocellin, R., and Velakoulis, D. (2007). Validity and reliability of the Behavioural assessment Tool for Cognition and Higher Function (BATCH) in neuropsychiatric patients. *Australian and New Zealand Journal of Psychiatry*, 41, 697–704. (RevSist)
- Miller, S. D., Duncan, B. L., Sorrell, R., & Brown, G. S. (2005). The partners for change outcome management system. *Journal of clinical psychology*, 61(2), 199–208. <https://doi.org/10.1002/jclp.20111>
- Milne, D., Reichelt, K., and Wood, E. (2001). Implementing HoNOS: An Eight Stage Approach. *Clinical Psychology and Psychotherapy*, 8, pp. 106–116. (RevSist)
- Ministerio de Sanidad y Consumo. (2007). *Estrategia en Salud Mental del Sistema Nacional de Salud 2006*. Madrid: Ministerio de Sanidad y Consumo.
- Ministerio de Sanidad y Política Social. (2010). *Evaluación de la Estrategia en Salud Mental del Sistema Nacional de Salud*. Ministerio de Sanidad, Política Social e Igualdad. https://www.msbs.gob.es/organizacion/sns/planCalidadSNS/pdf/excelencia/Resumen_Evaluacion.pdf
- Ministerio de Sanidad, Servicios Sociales e Igualdad. (2012). *Informe del Estado Actual de la Prevención en Salud Mental*. <https://consaludmental.org/publicaciones/Estadoactualprevencionsaludmental.pdf>
- Ministerio de Trabajo y Asuntos Sociales (IMSERSO). (2007). *Modelo de atención a las personas con Enfermedad Mental Grave*. Colección Documentos Serie Documentos Técnicos N. 21005
- Mokken, R.J. (1997). *Nonparametrics models for dichotomous responses*. En W.J. van der Linden y R.K. Hambleton (Eds.). *Handbook of Modern Item Response Theory*. New York: Springer.
- Mokkink, LB., Terwee, CB., Knol, DL., Stratford, PW., Alonso, J., Patrick, DL., Bouter, LM., and De Vet, HCW. (2010) The COSMIN checklist for evaluating the methodological quality of studies on measurement properties: A clarification of its content. *BMC Medical Research Methodology*, 10(22).
- Mokkink, LB., Terwee, CB., Patrick, DL., Alonso, J., Stratford, PW., Knol, DL., Bouter, LM., and De Vet, HCW. (2010). The COSMIN study reached international consensus on taxonomy, terminology, and

- definitions of measurement properties for health-related patient-reported outcomes. *Journal of Clinical Epidemiology*, 63, pp. 737-745. doi: 10.1016/j.jclinepi.2010.02.006
- Mokkink, LB., Terwee, CB., Patrick, DL., Alonso, J., Stratford, PW., Knol, DL., Bouter, LM., and De Vet, HCW. (2010). The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Quality of Life Research*, 19, pp. 539–549.
- Monk-Cunliffe, J., Kadra-Scalzo, G., Finamore, C., Dale, O., Khondoker, M., Barrett, B., ... Moran, P. (2023). Defining severity of personality disorder using electronic health records: short report. *BJPsych Open*, 9(5), e137. doi:10.1192/bjo.2023.509 **(RevSist)**
- Moses, L. (1995). Measuring effects without randomized trials? Options, problems, challenges. *Medical Care*, 33, AS8-AS14.
- Mötteli, S., Adamus, C., Deb, T., Fröbel, R., Siemerker, J., Richter, D. and Jäger, M. (2022). Independent Supported Housing for Non-homeless People with Serious Mental Illness: A Pragmatic Randomized Controlled Trial. *Frontiers in Psychiatry*, 12, 798275. doi: 10.3389/fpsy.2021.798275 **(RevSist)**
- Mötteli, S., Schori, D., Schmidt, H., Seifritz, E., and Jäger. (2018). Utilization and Effectiveness of Home Treatment for People with Acute Severe Mental Illness: A Propensity-Score Matching Analysis of 19 Months of Observation. *Frontiers in Psychiatry*, (9)495. **(RevSist)**
- Muehsam, J.P. (2019). Association between clinical observations and a mobile crisis team's level of care recommendations. *Community Mental Health Journal*, 55, 394-400. <https://doi.org/10.1007/s10597-018-0296-y> **(RevSist)**
- Mukherjee, H. and Sazhin, V. (2022). Predictors of functioning and clinical outcomes in inpatient with schizophrenia on clozapine augmented with antipsychotics. *Australasian psychiatry*, 29(6), pp. 655 – 658. DOI: 10.1177/10398562211037339 **(RevSist)**
- Mulder, C.L., Staring, A.B.P., Loos, J., Buwalda, V.J.A., Kuijpers, D., Sytema, S., and Wierdsma A.I. (2004). De Health of the Nation Outcome Scales (honos) als instrument voor 'routine outcome assessment'. *Tijdschrift Voor Psychiatrie*, 46(5), pp. 273-284. **(RevSist)**
- Mulligan, O., Helmi, L., Adamis, D., O'Mahony, E., McCarthy, G., and MacDonald, A. (2016). *Implementation of Health of the Nation Outcome Scale (HoNOS) in Outpatient Clinic, Sligo Mental Health Services: Feasibility and Agreement with Global Assessment of Functioning Scale*. 24th European Congress of Psychiatry/European Psychiatry 33S, EV785. <http://dx.doi.org/10.1016/j.eurpsy.2016.01.1770> **(RevSist)**
- Muncer, S., Bass, M., and Dawkin, M. (2016). Mokken analysis of the Health of the Nation Outcome Scales in acute inpatient and community samples. *Australasian Psychiatry*, 24(5), pp. 459-461. **(RevSist)**
- Murugesan, G., Amey, C.G., Deane, F.P., Jeffrey, R., Kelly, B., and Stain, H. (2007). Inpatient psychosocial rehabilitation in rural NSW: assessment of clinically significant change for people with severe mental illness. *Australian and New Zealand Journal of Psychiatry*, 41, pp. 343-350. **(RevSist)**

N

- Nash, L., Gorrell, J., Cornish, A., Rosen, A., Miller, V., and Tennant. (2004). Clinical outcome of an early psychosis intervention program: evaluation in a real-world context. *Australian and New Zealand Journal of Psychiatry*, 38, pp. 694–701. **(RevSist)**
- National Institute for Mental Health in England (2008). *Outcomes Compendium*. Birmingham: National Institute for Mental Health in England (NIMHE).
- National Institute for Mental Health. (2008). *Mental health Outcomes Compendium*. Birmingham: Autor.
- New South Wales Health. (2006). Mental Health Outcomes & Assessment Tools (MH-OAT) data collection reporting requirement 1 July 2006. *North Sydney: NSW Health*, 1–10.
- Newnham, E.A., Harwood, K.E., and Page, A.C. (2007). Evaluating the clinical significance of responses by psychiatric inpatients to the mental health subscales of the SF-36. *Journal of Affective Disorders*, 98, pp. 91–97. **(RevSist)**
- Newnham, E.A., Harwood, K.E., and Page, A.C. (2009). The subscale structure and clinical utility of the Health of the Nation Outcome Scale. *Journal of Mental Health*, 18(4), 326–334. <https://doi.org/10.1080/09638230802522486>
- Newnham, E.A., Harwood, K.E., and Page, A.C. (2009). The subscale structure and clinical utility of the Health of the Nation Outcome Scale. *Journal of Mental Health*, 18(4), pp. 326–334. **(RevSist)**
- Noll, S., Haag, S., Guidon, R. and Hölzer, S. (2023). A new case-mix based payment system for the psychiatric day care sector in Switzerland: proposed methods for developing the tariff structure. *Health Policy*, 131. <https://doi.org/10.1016/j.healthpol.2023.104797> **(RevSist)**
- Noordraven, E.L., Wierdsma, A.I., Blanken, P., Bloemendaal, A.F.T., and Mulder, C.L. (2018). The effect of financial incentives on patients' motivation for treatment: results of "Money for Medication," a randomized controlled trial. *BMC Psychiatry*, 18(144). **(RevSist)**
- Noorthoorn E.O., Smits, H. J. H., Penterman, E. J. M., Seelen-de Lang, B. L., Nieuwenhuis, J. G., Nijman, H. L. I. (2021). The associations of quality of life and general functioning with trauma, borderline intellectual functioning and mild intellectual disability in outpatients with serious mental illness. *Research in Developmental Disabilities*, 115. <https://doi.org/10.1016/j.ridd.2021.103988> **(RevSist)**
- Nordbrandt, M. S., Sonne, C., Mortensen, E. L. and Carlsson, J. (2020). Trauma-affected refugees treated with basic body awareness therapy or mixed physical activity as augmentation to treatment as usual. A pragmatic randomized controlled trial. *PLoS ONE* 15(3), e0230300. <https://doi.org/10.1371/journal.pone.0230300> **(RevSist)**
- Norma UNE66175. (2003). *Guía para la implantación de Sistemas de Indicadores*. AENOR Ediciones
- Nunnally, J. C. and Bernstein, I. H. (1994). *Psychometric Theory*. 3 Ed. New York: McGraw-Hill.

O

- O'Brien, S., McFarland, J., Kealy, B., Pullala, A., Saunders, J., Cullen, W., and Meagher, D. (2012). A randomized- controlled trial of intensive case management emphasizing the recovery model among patients with severe and enduring mental illness. *The Irish Journal of Psychological Medicine*, 181, 301–308. Doi: 10.1007/s11845-011-0795-0 **(RevSist)**

- O'Connell, J., Sunwoo, M., McGorry, P. and O'Donoghue, B. (2019). Characteristics and outcomes of young people with substance induced psychotic disorder. *Schizophrenia Research*, 206, pp. 257-262. <https://doi.org/10.1016/j.schres.2018.11.007> (RevSist)
- O'Connor, J. and Katyal, R. (2021). Perinatal mental health outcome measures in a mother and baby unit. *BJPsych Open*, 18(7) (Suppl 1), S96–S97. 10.1192/bjo.2021.290. (RevSist)
- Oficina Regional de la OMS para Europa. (2010). *User empowerment in mental health: a statement by the WHO Regional Office for Europe. Empowerment is not a destination, but a journey*. Copenhagen. <https://apps.who.int/iris/handle/10665/107275>
- Øiesvold, T., Bakkejord, T., and Sexton, JA. (2011). Concurrent validity of the Health of the Nation Outcome Scales compared with a patient-derived measure, the Symptom Checklist-90-Revised in out-patient clinics. *Psychiatry Research*, 187, pp. 297-300. (RevSist)
- Organización Mundial de la Salud. (2003). *Clasificación Estadística Internacional de las Enfermedades y Problemas Relacionados con la Salud (Décima Revisión)*. Washington: Organización Panamericana de la Salud.
- Organización Panamericana de la Salud (OPS). (2023). *Informe mundial sobre la salud mental: Transformar la salud mental para todos*. Washington, D.C. <https://doi.org/10.37774/9789275327715>.
- Orrell, M., Yard, P., Handysides, J. and Schapira, R. (1999). Validity and reliability of the Nation Outcome scales in psychiatric patients in the community. *British Journal of Psychiatry*, 174, pp. 409-412. (RevSist)
- Osta-Samanes, R., García-Tejedor, A., & Muro, C. (2012). Integración social del anciano institucionalizado. *Gerokomos*, 23(4), 172-176. <https://doi.org/10.4321/S1134-928X2012000400006> (RevSist)

P

- Page, A.C., Hooke, G. R., Rutherford, E. M. (2001). Measuring mental health outcomes in a private psychiatric clinic: Health of the Nation Outcome Scales and Medical Outcomes Short Form SF-36. *Australian and New Zealand Journal of Psychiatry*, 35, pp. 377–381. (RevSist)
- Parabiaghi, A., Barbato, A., D'Avanzo, B., Erlicher, A., & Lora, A. (2005). Assessing reliable and clinically significant change on HoNOS: A method for displaying longitudinal data. *Australasian and New Zealand Journal of Psychiatry*, 39, 719-725. <https://doi.org/10.1080/j.1440-1614.2005.01656.x>
- Parabiaghi, A., Barbato, A., D'Avanzo, B., Erlicher, A., and Lora, A. (2005). Assessing reliable and clinically significant change on Health of the Nation Outcome Scales: method for displaying longitudinal data. *Australian and New Zealand Journal of Psychiatry*, 39, pp. 719-725. (RevSist)
- Parabiaghi, A., Kortrijk, H.E., & Mulder, C.L. (2014). Defining multiple criteria for meaningful outcome in routine outcome measurement using the Health of the Nation Outcome Scales. *Social Psychiatry and Psychiatric Epidemiology*, 49(2), 291-305. <https://doi.org/10.1007/s00127-013-0750-7>
- Parabiaghi, A., Rapisarda, F., D'Avanzo, B., Erlicher, A., Lora, A., and Barbato, A. (2011). Measuring clinical change in routine mental health care: differences between first time and longer-term service

- users. *Australian and New Zealand Journal of Psychiatry*, 45, pp. 558 – 568. DOI: 10.3109/00048674.2011.580450 (RevSist)
- Parker, S., Arnautovska, U., McKeon, G. and Kisely, S. (2021). The association between discontinuation of community treatment orders and outcomes in the 12-months following discharge from residential mental health rehabilitation. *International Journal of Law and Psychiatry*, 74, 101664. <https://doi.org/10.1016/j.ijlp.2020.101664>. (RevSist)
- Parker, S., Arnautovska, U., Siskind, D., Dark, F., McKeon, G., Korman, N., and Harris, M. (2020). Community-care unit model of residential mental health rehabilitation services in Queensland, Australia: predicting outcomes of consumers 1-year post discharge. *Epidemiology and Psychiatric Sciences*, 29, e109. doi:10.1017/S2045796020000207 (RevSist)
- Parker, S., Dark, F., Newman, E., Hanley, D., McKinlay, W. and Meurk, C. (2019). Consumers' understanding and expectations of a community-based recovery-oriented mental health rehabilitation unit: a pragmatic grounded theory analysis. *Epidemiol Psychiatr Sci*, 28(4), pp. 408-417. doi: 10.1017/S2045796017000749 (RevSist)
- Parker, S., Dark, F., Newman, E., Korman, N., Meurk, C., Siskind, D., and Harris, M. (2016)., longitudinal comparative evaluation of the equivalence of an integrated peer-support and clinical staffing model for residential mental health rehabilitation: a mixed methods protocol incorporating multiple stakeholder perspectives. *Biomed Central Psychiatry*, 16(179). (RevSist)
- Parker, S., Siskind, D., Hermens, D. F., Dark, F., McKeon, G., Korman, N., Arnautovska, U., Harris, M. and Whiteford, H. (2019). A Comprehensive Cohort Description and Statistical Grouping of Community-Based Residential Rehabilitation Service Users in Australia. *Frontiers in Psychiatry* 10, 798. doi: 10.3389/fpsyt.2019.00798 (RevSist)
- Páv, M., Vnřuková, M. and Sebalo, I. (2022). Factors Affecting Length of Inpatient Forensic Stay: Retrospective Study from Czechia *Front. Psychiatry*, 13, 825615. doi: 10.3389/fpsyt.2022.825615 (RevSist)
- Pec, O., Bob, P., Pec, J., and Hrubcova., A. (2017). Psychodynamic day treatment programme for patients with schizophrenia spectrum disorders: Dynamics and predictors of therapeutic change. *Psychology and Psychotherapy: Theory, Research and Practice*. DOI:10.1111/papt.12153 (RevSist)
- Pěč, O., Čechová, D., Pěčová, J., Dragomirecká, E., Šelepová, P., Probstová, V., and Bubela, M. (2009). HoNOS (Hodnocení zdravotního stavu) – adaptace nástroje na hodnocení symptom a sociálních funkcí u závažně duševně nemocných v českých podmínkách a jeho použití. *Česká a slovenská psychiatrie*, 105(6–8), 245–249. (RevSist)
- Pedreira Massa, J.L., Eguiagaray García. M. (1996). Evaluación de un Servicio de Psiquiatría Infantil a través de un registro acumulativo de casos psiquiátricos. *Rev Psiquiatr Infant Juv*, 4, 263-273.
- Pedrini, L., Lanfredi, M., Ferrari, C. et al. (2018). Development and Validation of the Health of the Nation Outcome Scales-Residential Facility (HoNOS-RF). *Psychiatr Q*, 89, 461–473. <https://doi.org/10.1007/s11126-017-9548-1>
- Pelizza, L., Maestri, D., Leuci, E., Paulillo, G., Ceroni, P., Quattrone, E., and Pellegrini, P. (2022). Personal Health Budget: a new rehabilitation approach for severe mental illness within a caring community. *European Psychiatry*, 65(S1), S246–S246. doi:10.1192/j.eurpsy.2022.635 (RevSist)
- Penington, E., Lennox, B., Geulayov, G., Hawton, K. and Tsiachristas, A. (2022). The early impact of the COVID-19 pandemic on patients with severe mental illness: An interrupted time-series study in

- South-East England. *European Psychiatry*, 65(1), e31, 1–10, <https://doi.org/10.1192/j.eurpsy.2022.22> (RevSist)
- Penno, S. J., Hamilton, B. and Petrakis, M. (2017). Early Intervention in Psychosis: Health of the Nation Outcome Scales (HoNOS) Outcomes from a Five-Year Prospective Study. *Archives of Psychiatric Nursing*, 31(6), 553-560. <https://doi.org/10.1016/j.apnu.2017.07.003>. (RevSist)
- Peritogiannis, V., Grammeniati, A., Gogou, A., Gioti, P. and Mavreas, V. (2019). Very Long-Term Outcome of Community-Dwelling Patients with Schizophrenia Spectrum Disorders in Rural Greece. *J Nerv Ment Dis*, 207(12), 1012-1018. doi: 10.1097/NMD.0000000000001068. (RevSist)
- Phuaphanprasert, B., Srisurapanont, M., Silpakit, C., Pannarunothai, S., Udomratn, P., Geater, A., & Prapaphomrn, P. (2007). Reliability and validity of the Thai version of the Health of the Nation Outcome Scales (hoNOS). *Journal of the Medical Association of Thailand = Chotmaihet thangphaet*, 90(11), 2487–2493.
- Pirkis, J., Burgess, P., Coombs, T. et al. (2005). Routine measurement of outcomes in Australia's public sector mental health services. *Aust N Z Health Policy* 2(8). <https://doi.org/10.1186/1743-8462-2-8>
- Pirkis, J.E., Burgess, P.M., Kirk, P.K., Dodson, S., Coombs, T.J., & Williamson, M.K. (2005). A review of the psychometric properties of the Health of the Nation Outcome Scales (HoNOS) family of measures. *Health and Quality of Life Outcomes*, 3, 76, 1-12. <https://doi.org/10.1186/1477-7525-3-76>
- Pons, I., Antón, I., et al. (1992). Los registros acumulativos de casos psiquiátricos como sistema de información y evaluación dentro del marco de la reforma psiquiátrica. *Rev Asoc Esp Neuropsiquiatría*, 12, 115-124.
- Prabhu, R., & Oakley-Browne, M. (2008). The use of the Health of the Nation Outcome Scale in an outreach rehabilitation program. *Australasian Psychiatry*, 16(3), 195-199. <https://doi.org/10.1080/10398560701784813>
- Prabhu, R., and Oakley Browne, M. (2008). The use of the Health of the Nation Outcome Scale in an outreach rehabilitation program. *Australasian Psychiatry*, 16(3), pp. 195-199. (RevSist)
- Preston, N.J. (2000). The Health of the Nation Outcome Scales: Validating factorial structure and invariance across two health services. *Australasian and New Zealand Journal of Psychiatry*, 34, 512-519. <https://doi.org/10.1046/j.1440-1614.2000.00726.x>
- Preston, N.J. (2000). The Health of the Nation Outcome Scales: validating factorial structure and invariance across two health services. *Australian and New Zealand Journal of Psychiatry*, 34, pp. 512-519. (RevSist)
- Preti, A., Pisano, A., Cascio, M.T., Galvan, F., Monzani, E., Meneghelli, A., and Cocchi, A. (2012). Validation of the Health of the Nation Outcome Scales as a routine measure of Outcome in early intervention programmes. *Early Intervention in Psychiatry*, 6, pp. 423-431. (RevSist)
- Prinsen, C. A. C., Mokkink, L. B., Bouter, L. M., Alonso, J., Patrick, D. L., de Vet, H. C. W., & Terwee, C. B. (2018). COSMIN guideline for systematic reviews of patient-reported outcome measures. *Quality of life research: an international journal of quality of life aspects of treatment, care and rehabilitation*, 27(5), 1147–1157. <https://doi.org/10.1007/s11136-018-1798-3>
- Puschner, B., Becker, T., & Bauer, S. (2015). Routine outcome measures in Germany. *International Review of Psychiatry*, 27(4), 329–337. <https://doi.org/10.3109/09540261.2015.1014025>

R

- Red de Salud Mental de Bizkaia. (2019). *Unidad Docente Multiprofesional de la Red de Salud Mental de Bizkaia (UDM-RSMB). Presentación UDM-RSMB.* https://www.osakidetza.euskadi.eus/contenidos/informacion/rsmb_unidad_docente_multiprof/es_rsmb/adjuntos/2019%2001%2030%20-%2001%20Presentaci%C3%B3n%20RSMB.pdf
- Redlich Bossy, M., Mueller, D., Seifritz, E., Vetter, S. and Egger, S. T. (2020). Feasibility and Efficacy of a Psychological Therapy for Patients with a Schizophrenic Psychosis in an Inpatient Setting: Study Protocol of a Randomized Switch Controlled Trial. *Frontiers in Public Health*, 8:391. doi: 10.3389/fpubh.2020.00391 (RevSist)
- Rees, A., Richards, A., and Shapiro, DA. (2004). Utility of the HoNOS in measuring change in a Community Mental Health Care population. *Journal of Mental Health*, 13(3), pp. 295-304. (RevSist)
- Reise, S.P., Bonifay, W.E., & Haviland, M.G. (2013). Scoring and modeling psychological measures in the presence of multidimensionality. *Journal of Personality Assessment*, 95(2), 129-140. <https://doi.org/10.1080/00223891.2012.725437>
- Ribas, S., Infante, P., and Fernández-González, D. (2016). Rehabilitation programme of Residential Homes: compare the results in two different models. *International Journal of Integrated Care*, 16(6), A43, pp. 1-8. DOI: <http://doi.org/10.5334/ijic.2986> (RevSist)
- Richardson, T., Dasyam, B., Courtney, H., White, L., Tedbury, J., Butt, J. and Newman-Taylor, K. (2019). Predictors of disengagement from cognitive behavioural therapy for psychosis in a National Health Service setting: A retrospective evaluation. *British Journal of Clinical Psychology*, 58(4), pp. 440-451. DOI:10.1111/bjc.12222 (RevSist)
- Rock, D. and Preston, N. (2001). HoNOS: is there any point in training clinicians? *Journal of Psychiatric and Mental Health Nursing*, 8, pp. 405–409. (RevSist)
- Rodriguez, A., & Reise, S.P. (2016). Evaluating Bifactor Models: Calculating and interpreting statistical indices. *Psychological Methods*, 21(2), 137-150. <https://doi.org/10.1037/met0000045>
- Rosseel, Y. (2018). *Lavaan: Latent variable analysis*. <https://CRAN.R-project.org/package=lavaan>
- Roy, A., Matthews, H., Clifford, P., Fowler, V., & Martin, D. M. (2002). Health of the Nation Outcome Scales for People with Learning Disabilities (HoNOS-LD). *The British journal of psychiatry: the journal of mental science*, 180, 61–66. <https://doi.org/10.1192/bjp.180.1.61>

S

- Sacristán, J.A., Galende, I., Soto, J. (1998). Estudios naturalísticos para valorar la efectividad de los medicamentos tras su comercialización: ¿por qué, cuándo y cómo? *Atención Primaria*, 22(3), 182- - 185.
- Salvi, G., Leese, M., and Slade, M. (2005). Routine use of mental health outcome assessments: choosing the measure. *British Journal of Psychiatry*, 186, pp. 146-152. (RevSist)
- Samejima, F. (1969). Estimation of latent ability using a response pattern of graded scores. *Psychometrika Monograph Supplement*, 34, 100–114.

- Sanches, S.A., Swildens, W.E., van Busschbach, J.T. et al. (2019). Identifying social participation subgroups of individuals with severe mental illnesses: a latent class analysis. *Social Psychiatry and Psychiatric Epidemiology*, 54, 1067–1077. <https://doi.org/10.1007/s00127-019-01704-y> (RevSist)
- Schallock, R. L. (1995). *Outcome-Based Evaluation*. Plenum Press: New York.
- Schmitter-Edgecombe, M., McAlister, C., & Weakley, A. (2012). Naturalistic assessment of everyday functioning in individuals with mild cognitive impairment: The day-out task. *Neuropsychology*, 5, 631-641. <https://doi.org/10.1037/a0029352>
- Schoenwald, S.K., Garland, A.F., Chapman, J.E., Frazier, S.L., Sheidow, A.J., & Southam-Gerow, M.A. (2011). Toward the effective and efficient measurement of implementation fidelity. *American Policy in Mental Health*, 38(1), 32-43.
- Scholtes, V.A., Terwee, C.B., & Poolman, R.W. (2001). What makes a measurement instrument valid and reliable? *Injury*, 42(3), 236-240. <https://doi.org/10.1016/j.injury.2010.11.042>
- Schroeder, S. A. (1987). Outcome assessment 70 years later: are we ready? *The New England journal of medicine*, 316(3), 160–162. <https://doi.org/10.1056/NEJM198701153160309>
- Segal, S. P., Rimes, L., and Hayes, S. L. (2019). The utility of outpatient commitment: Reduced-risks of victimization and crime perpetration. *European Psychiatry*, 56(1), pp. 97–104. doi:10.1016/j.eurpsy.2018.12.001 (RevSist)
- Sekharan, L., Jagadheesan, K., Das, P., Lakra, V., West, K. and Baldwin, M. (2021). Short-term and 1-year outcome of patients' with borderline personality admitted to a short-term recovery-oriented residential service. *Australasian Psychiatry*, 29(6), pp. 581-585. doi:10.1177/1039856220961646 (RevSist)
- Seow, L. L. Y., Page, A. C. and Hooke, G. R. (2020). Severity of borderline personality disorder symptoms as a moderator of the association between the use of dialectical behaviour therapy skills and treatment outcomes. *Psychotherapy Research*, DOI: 10.1080/10503307.2020.1720931 (RevSist)
- Shamseer, L., Moher, D., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., Shekelle, P., Stewart, LA. The PRISMA-P Group. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation. *BMJ Open*. doi: 10.1136/bmj.g7647
- Sharma, V.K., Wilkinson, G., and Fear, S. (1999). Health of the Nation Outcome Scales: a case study in general psychiatry. *British Journal of Psychiatry*, 174, pp. 391-398. (RevSist)
- Shergill, S.S., Shankar, K.K., Seneviratna, K., Orrell, M.W. (1999). The validity and reliability of the Health of the Nation Outcome Scales (HoNOS) in the elderly. *Journal of Mental Health (UK)*, 8, 511-521.
- Sherrill, J. (2015). *Pragmatic Strategies for Assessing Psychotherapy Quality in Practice*. <https://www.nimh.nih.gov/funding/grant-writing-and-application-process/concept-clearances/2015/pragmatic-strategies-for-assessing-psychotherapy-quality-in-practice>
- Siskind, D., Harris, M., Diminic, S., Carstensen, G., Robinson, G., and Whiteford, H. (2014). Predictors of mental health-related acute service utilisation and treatment costs in the 12 months following an acute psychiatric admission. *Australian & New Zealand Journal of Psychiatry*, 48(11), pp. 1048–1058. DOI: 10.1177/0004867414543566 (RevSist)

- Siskind, D., Harris, M., Kisely, S., Brogan, J., Pirkis, J., Crompton, D., and Whiteford, H. (2013). A retrospective quasi-experimental study of a community crisis house for patients with severe and persistent mental illness. *Australian & New Zealand Journal of Psychiatry*, 47(7), pp. 667–675. DOI: 10.1177/0004867413484369 (RevSist)
- Skar-Fröding, R., Clausen, H., Šaltytė Benth, J., Ruud, T., Slade, M. and Heiervang, K. S. (2022). Associations between personal recovery and service user-rated versus clinician-rated clinical recovery, a cross-sectional study. *BMC Psychiatry*, 22(1), 42. doi: 10.1186/s12888-022-03691-y (RevSist)
- Slade, M. (2002). What outcomes to measure in routine mental health services, and how to assess them: a systematic review. *The Australian and New Zealand journal of psychiatry*, 36(6), 743–753. <https://doi.org/10.1046/j.1440-1614.2002.01099.x>
- Slade, M., Beck, A., Bindman, J., Thornicroft, G., and Wright, S. (1999). Routine clinical outcome measures for patients with severe mental illness: CANSAS and HoNOS. *British Journal of Psychiatry*, 174, pp. 404-408. (RevSist)
- Slade, M., Cahill, S., Kelsey, W., Powell, R., and Strathdee, G. (2002). Threshold 2: the reliability, validity and sensitivity to change of the Threshold Assessment Grid (TAG). *Acta Psychiatrica Scandinavica*, 106, pp. 453-460. (RevSist)
- Slade, M., Thornicroft, G. and Glover, G. (1999). The feasibility of routine outcome measures in mental health. *Social Psychiatry and Psychiatric Epidemiology*, 34, 243–249.
- Smith, M., Baxendine, S., Dempster, F. and Fitzpatrick, S. (2023). *HoNOSI Guide for New Zealand Clinicians HoNOSI | Guide for New Zealand Clinicians 2*. 10.13140/RG.2.2.30098.63683.
- Smith, W., Vitoratou, S., McCrone, P., and Patel, A. (2017). Exploring the factor structure of the Health of the Nation Outcome Scale (HoNOS) in a sample of patients with schizophrenia, schizotypal and delusional disorders. *Health and Quality of Life Outcomes*, 15, pp. 141-150. (RevSist)
- Soni, N., Roberts, S. and Branjerdporn, G. (2022). Exploring Discharge Outcomes and Readmission Rates of Mothers Admitted to a Psychiatric Mother and Baby Unit. *Psychiatric Quarterly*, 93, 393–407. <https://doi.org/10.1007/s11126-021-09956-1> (RevSist)
- Speak, B. L., & Muncer, S. J. (2016). Factorial structure of the Health of the Nation Outcome Scales: An ordinal confirmatory factor analysis using a national sample of clinician ratings in England. *International Journal of Mental Health Nursing*, 25, 87-98. <https://doi.org/10.1111/inm.12199> (RevSist)
- Speak, B., & Muncer, S. (2015). The structure and reliability of the Health of the Nation Outcomes Scale. *Australasian Psychiatry*, 23(1), 66-68. <https://doi.org/10.1177/1039856214563851>
- Speak, B., and Muncer, S. (2015). The structure and reliability of the Helath of the Nation Outcome Scales. *Australasian Psychiatry*, 23(1), pp. 66-68. (RevSist)
- Sportel, B. E., Aardema, H. and Boonstra, N. et al. (2023). Measuring recovery in participants with a schizophrenia spectrum disorder: validation of the Individual Recovery Outcomes Counter (I.ROC). *BMC Psychiatry* 23(296). <https://doi.org/10.1186/s12888-023-04763-3> (RevSist)
- Stedman, T., Yellowlees, P., Drake, S., Chant, D., Clarke, R., and Chapple, B. (2000). The perceived utility of six selected measures of consumer outcomes proposed for routine use in Australian mental health services. *Australian and New Zealand Journal of Psychiatry*, 34, pp. 842–849. (RevSist)

- Stedman, T., Yellowlees, P., Drake, S., Chant, D., Clarke, R., Chapple B. (2000). The Perceived Utility of Six Selected Measures of Consumer Outcomes Proposed for Routine Use in Australian Mental Health Services. *Australian & New Zealand Journal of Psychiatry*, 34(5), 842-849. doi:10.1080/j.1440-1614.2000.00803.x
- Steel, Z., McDonald, R., Silove, D., Bauman, A., Sandford, P., Herron, J., and Minas, I.H. (2006). Pathways to the first contact with specialist mental health care. *Australian and New Zealand Journal of Psychiatry*, 40, pp. 347–354. **(RevSist)**
- Stephenson, LA., Macdonald, AJD., Seneviratne, G., Waites, F., and Pawlby, S. (2018). Mother and Baby Units matter: improved outcomes for both. *The British Journal of Psychiatry*, 4, pp. 119-125. doi: 10.1192/bjo.2018.7 **(RevSist)**
- Stewart, M. (2009). Service user and significant other versions of the Health of the Nation Outcome Scales. *Australasian Psychiatry*, 17(2). doi: 10.1080/10398560802596116 **(RevSist)**
- Stewart, M. (2009). Service user and significant other versions of the Health of the Nation Outcome Scales. *Australasian Psychiatry*, 17, 156–63.
- Stiekema, A.P.M., Looijmans, A., Van Der Meer, L., Bruggeman, R., Schoevers, R.A., Corpeleijn, E., and Jörg, F., (2018). Effects of a lifestyle intervention on psychosocial well-being of severe mentally ill residential patients: ELIPS, a cluster randomized controlled pragmatic trial. *Schizophr Res.*, 199, 407-413. doi: 10.1016/j.schres.2018.02.053. **(RevSist)**
- Strømme, M. F., Bartz-Johannessen, C., Kroken, R. A., Mehlum, L. and Johnsen, E. (2022). Overactive, aggressive, disruptive and agitated behavior associated with the use of psychotropic medications in schizophrenia. *Schizophrenia Research*, 248, 35-41. <https://doi.org/10.1016/j.schres.2022.07.014> **(RevSist)**
- Stubbs, B., Mueller, C., Gaughran, F., Lally, J., Vancampfort, D., Lamb, S. E., Koyanagi, A., Sharma, S., Stewart, R. and Perera, G. (2018). Predictors of falls and fractures leading to hospitalization in people with schizophrenia spectrum disorder: A large representative cohort study. *Schizophrenia Research*, 201. Pp. 70-78. <https://doi.org/10.1016/j.schres.2018.05.010>. **(RevSist)**
- Sugarman, P., Walker, L., and Dickens, G. (2009). Managing outcome performance in mental health using HoNOS: experience at St Andrew's Healthcare. *Psychiatric Bulletin* 33, pp. 285-288. doi: 10.1192/pb.bp.108.022475 **(RevSist)**
- Sugarman, P., y Walker, L. (2007). *Health of the Nation Outcome Scale for Users of Secure and Forensic Services. How to Use HoNOS-Secure Version 2b*. https://www.rcpsych.ac.uk/docs/default-source/improving-care/ccqi/honos/honos-secure-v2b-explanation.pdf?sfvrsn=9fedeab0_2
- Sunwoo, M., O'Connell, J., Brown, E., Lin, A., Wood, S. J., McGorry, P. and O'Donoghue, B. (2020). Prevalence and outcomes of young people with concurrent autism spectrum disorder and first episode of psychosis. *Schizophrenia Research*, 216, 310-315. <https://doi.org/10.1016/j.schres.2019.11.037>. **(RevSist)**
- Sweeney, J., Adamis, D., Helmi, L. et al. (2019). Implementation of outcome measurement (HoNOS) in an outpatient psychiatric clinic in Sligo/Leitrim mental health service. *Irish Journal of Medical Science*, 188, 1329–1335. <https://doi.org/10.1007/s11845-019-02015-6> **(RevSist)**
- Swildens, W. E., Visser, E., Bähler, M., Bruggeman, R., Delespaul, P., van der Gaag, M., de Haan, L., Keet, R., Nijssen, Y., van Os, J., Pijnenborg, G. M., Slooff, C., de Vos, A., van Weeghel, J., Wunderink, L., Mulder, C. L., and Wiersma, D. (2018). Functional recovery of individuals with

serious mental illnesses: Development and testing of a new short instrument for routine outcome monitoring. *Psychiatric Rehabilitation Journal*, 41(4), 341–350. <https://doi.org/10.1037/prj0000320>
(RevSist)

T

Taylor, C. L., Stewart, R. J. and Howard, L. M. (2019). Relapse in the first three months postpartum in women with history of serious mental illness. *Schizophrenia Research*, 204, 46-54. <https://doi.org/10.1016/j.schres.2018.07.037>.
(RevSist)

Taylor, K. M., Orucu, E., Nandha, S. and Cella, M. (2023). Psychosis Early Intervention Across the Life Span: A Service Perspective. *Community Mental Health Journal*, 59, pp. 266–274. <https://doi.org/10.1007/s10597-022-01004-x>
(RevSist)

Te Pou (2021). *HoNOS Guide for New Zealand Clinicians*. <https://www.tepou.co.nz/resources/honos-guide-for-new-zealand-clinicians>

Tenneij, N., Didden, R., Veltramp, E., and Koot, HM. (2009). Reliability and validity of the HoNOS-LD and HoNOS in a sample of individuals with mild to borderline intellectual disability and severe emotional and behavior disorders. *Journal of Mental health Research in Intellectual Disabilities*, 2, pp. 188-200.
(RevSist)

Terwee, C. B., Prinsen, C. A. C., Chiarotto, A., Westerman, M. J., Patrick, D. L., Alonso, J., Bouter, L. M., de Vet, H. C. W., & Mokkink, L. B. (2018). COSMIN methodology for evaluating the content validity of patient-reported outcome measures: a Delphi study. *Quality of life research: an international journal of quality of life aspects of treatment, care and rehabilitation*, 27(5), 1159–1170. <https://doi.org/10.1007/s11136-018-1829-0>

Terwee, CB., Bot, SDM., De Boer, MR., Van der Windt, DAWM., Knol, DL., Dekker, J., Bouter, LM., and De Vet, HCW. (2007). Quality criteria were proposed for measurement properties of health status questionnaires. *Journal of Clinical Epidemiology*, 60, pp. 34-42.

Terwee, CB., Mokkink, LB., Knol, DL., Ostelo, RWJG., Bouter, LM., and De Vet, HCW. (2012). Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. *Quality of Life Research*, 21, pp. 651–657. DOI 10.1007/s11136-011-9960-1

The jamovi Project. (2020). *Jamovi (Version 1.2) [Computer Software]*. Retrieved from <https://www.jamovi.org>

Theodoridou, A., Jaeger, M., Ketteler, D., Kawohl, W., Lauber, C., Hoff, P., and Rössler, W. (2011). The concurrent validity and sensitivity of change of the german version of the Health of the Nation Outcome Scales in a psychiatric patient setting. *Psychopathology*, 44, pp. 391-397.
(RevSist)

Trauer, T. (1998). The Health of the Nation Outcome Scales in Outcome Measurement: A Critical Review. *Australasian Psychiatry*, 6(1), 11-14. doi:10.3109/10398569809082301

Trauer, T. (1999). The subscale structure of the Health of the Nation Outcome Scales (HoNOS). *Journal of Mental Health*, 8(5), 499–509. <https://doi.org/10.1080/09638239917193>
(RevSist)

Trauer, T. (2004). Consumer and service determinants of completion of a consumer self-rating outcome measure. *Australasian Psychiatry*, 12(1), 48-54.
(RevSist)

- Trauer, T. (2004b). Use of a routine outcome measure in a consultation–liaison mental health service. *Australasian Psychiatry*, 12(2), pp. 139–144. **(RevSist)**
- Trauer, T. (2010). *Outcome Measurement in Mental Health: Theory and Practice*. Cambridge: Cambridge University Press.
- Trauer, T., and Nagel, T. (2012). Outcome measurement in adult Indigenous mental health consumers. *Advances in Mental Health*, 11(1), pp. 106–116. **(RevSist)**
- Trauer, T., Callaly, T., and Hantz, P. (1999). The measurement of improvement during hospitalisation for acute psychiatric illness. *Australian and New Zealand Journal of Psychiatry*, 33, pp. 379–384. **(RevSist)**
- Trauer, T., Callaly, T., Hantz, P., Little, J., Shields, R., and Smith, J. (1999). Health of the Nation Outcome Scales. Results of the Victorian field trial. *British Journal of Psychiatry*, 174, pp. 380–388. **(RevSist)**
- Tucker, S., Hughes, J., Jolley, D., Buck, D., Hargreaves, C. and Challis, D. (2018). Reconfiguring in-patient services for adults with mental health problems: changing the balance of care. *BJPsych Open*, 4, pp. 420–426. doi: 10.1192/bjo.2018.60 **(RevSist)**
- Tulloch, A. D., David, A. S. and Thornicroft, G. (2015). Exploring the predictors of early readmission to psychiatric hospital. *Epidemiology and Psychiatric Sciences*, pp. 1–13. DOI:10.1017/S2045796015000128 **(RevSist)**
- Twomey, C., Prina, AM., Baldwin, DS, Das-Munshi, J., Kingdon, D., Koeser, L., Prince, MJ., Stewart, R., Tulloch, AD., Cieza, A. (2016). Utility of the Health of the Nation Outcome Scales (HoNOS) in Predicting Mental Health Service Costs for Patients with Common Mental Health Problems: Historical Cohort Study. *PLoS ONE* 11(11): e0167103. <https://doi.org/10.1371/journal.pone.0167103> **(RevSist)**

U

- Uriarte, J. J., Iglesias, N., Penas, P., Moreno-Calvete, M. C., Álvarez-González, A. and Iraurgi, I. (2024). Evidencias de Adecuación Psicométrica de la Versión Española de la Escala HoNOS. *Revista Iberoamericana de Diagnóstico y Evaluación. RIDEP*, 72(2), 144–155. <https://doi.org/10.21865/RIDEP72.2.10> **(RevSist)**
- Uriarte, J., Beramendi, V., Medrano, J., Wing, J., Beevor, A., & Curtis, R. (1999). Presentación de la traducción al castellano de la Escala HoNoS (Health of the Nation Outcome Scales). *Psiquiatría Pública*, 11(4), 93–101.

V

- Valerie Moran, V., and Jacobs, R. (2017). Investigating the relationship between costs and outcomes for English mental health providers: a bi-variate multi-level regression analysis. *European Journal of Health Economics*, 19 (709). <https://doi.org/10.1007/s10198-017-0915-5> **(RevSist)**
- Van Kranenburg, G.D., van den Brink, R.H.S., Mulder, W.G. et al. (2019). Clinical effects and treatment outcomes of long-term compulsory in-patient treatment of treatment-resistant patients with severe mental illness and substance-use disorder. *BMC Psychiatry*, 19(270). <https://doi.org/10.1186/s12888-019-2254-9> **(RevSist)**

- Vargas, M.L., Quintanilla, P., Lozano, A., Sendra-Gutiérrez, J.M. (2013). Riesgo psicosocial asociado al envejecimiento en el trastorno mental grave: validación y baremación de la Escala de Riesgo Psicosocial FEAFES Castilla y León (ERP). *Rehabilitación Psicosocial*, 10(1), pp. 3-10. **(RevSist)**
- Vázquez-Bourgon, J., Gómez Ruiz, E., Hoyuela Zatón, F., Salvador Carulla, L., Ayasa Arriola, R., Tordesillas Gutiérrez, D. and Crespo Facorro, B. (2021). Diferencias en la efectividad clínica y funcional, entre trastornos psiquiátricos, de un hospital de día psiquiátrico de agudos para pacientes con enfermedad mental aguda. *Revista de Psiquiatría y Salud Mental*, 14(1), 40-49. <https://doi.org/10.1016/j.rpsm.2019.04.001>. **(RevSist)**
- Veereschild, H. M., Noorthoorn, E. O., Nijman, H. L. I., Mulder, C. L., Dankers, M., Van der Veen, J. A., Loonen, A. J. M. and Hutschemaekers, G. J. M. (2020). Diagnose, indicate, and treat severe mental illness (DITSMI) as appropriate care: A three-year follow-up study in long-term residential psychiatric patients on the effects of re-diagnosis on medication prescription, patient functioning, and hospital bed utilization. *European Psychiatry*, 63(1), e47, 1–8. <https://doi.org/10.1192/j.eurpsy.2020.46> **(RevSist)**
- Verheesen, S. M. H., ten Doesschate, F., van Schijndel, M. A. et al. (2021). Intoxicated persons showing challenging behavior demand complexity interventions: a pilot study at the interface of the ER and the complexity intervention unit. *European Archives of Psychiatry and Clinical Neuroscience*, 271, 903–913. <https://doi.org/10.1007/s00406-020-01162-7> **(RevSist)**
- Vita, A., Deste, G., Barlati, S., De Peri, L., Giambra, A., Poli, R., Keefe, RSE., and Sacchetti, E. (2013). Interview-based assessment of cognition in schizophrenia: applicability of the Schizophrenia Cognition Rating Scale (SCoRS) in different phases of illness and settings of care. *Schizophrenia Research*, 146, pp. 217-223. **(RevSist)**

W

- Wade, D., Harrigan, S., Harris, M.G., Edwards, J., and McGorry, P.D. (2006). Pattern and correlates of inpatient admission during the initial acute phase of first-episode psychosis. *Australian and New Zealand Journal of Psychiatry*, 40, pp. 429–436. **(RevSist)**
- Waghorn, G., Chant, D., and King, R. (2005). Work-related subjective experiences among community residents with schizophrenia or schizoaffective disorder. *Australian and New Zealand Journal of Psychiatry*, 39, pp. 288–299. **(RevSist)**
- Waikato Evaluation Team. (2005). *Functional Outcome Measures for Children and Youth, Adults and Older Adult Consumers in New Zealand Mental Health Services*. Hamilton, NZ: University of Waikato.
- Walton, M. K., Powers, J. H., Hobart, J., Patrick, D. L., Marquis, P., Vamvakas, S., Isaac, M., Molsen, E., Cano, S. J., & Burke, L. (2015). Clinical outcome assessments: Conceptual foundation–report of the ISPOR clinical outcomes assessment –emerging good practices for outcomes research task force. *Value Health*, 18(6), 741-752. <https://doi.org/10.1016/j.jval.2015.08.006>
- Whitehead, R., Hopkins, L., Hughes, E., Kehoe, M. and Pedwell, G. (2021). “Everyone on the Same Team, All Working Together”: Implementing a Co-ordinated Multidisciplinary Approach to Supporting Young People with Co-occurring Intellectual Disability and Mental Health Issues. *Journal of Mental Health Research in Intellectual Disabilities*, 14(1), 1-22. DOI: 10.1080/19315864.2020.1864689 **(RevSist)**

- Wild, D., Eremenco, S., Mear, I., Martin, M., Houchin, C., Gawlicki, M., Hareendran, A., Wiklund, I., Yee-Chong, L., von-Maltzahn, R., Cohen, L., & Molsen, E. (2009). Multinational trials – recommendations on the translations required, approaches to using the same language in different countries, and the approaches to support pooling the data: the ISPOR patient-reported outcomes translation & linguistic validation good research practices task force report. *Value Health*, 12, 430-440. <https://doi.org/10.1111/j.1524-4733.2008.00471.x>
- Wilkinson, L., & The Task Force on Statistical Inference APA Board of Scientific. (1999). Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist*, 54(8), 594–604. <https://doi.org/10.1037//0003-066x.54.8.594>
- Williams, B., Speak B., Hay, P., and Muncer, SJ. (2014). An evaluation of the independence of the Health of the Nation Outcome Scales. *Australasian Psychiatry*, 22(5), pp. 473-475. **(RevSist)**
- Wing, J. K., Beevor, A. S., Curtis, R. H., Park, S.B.G., Hadden, S., & Burns, A. (1998). Health of the Nation Outcome Scales (HoNOS). *British Journal of Psychiatry*, 172, 11-18. <https://doi.org/10.1192/bjp.172.1.11>
- Wing, J.K., Curtis, R.H. & Beevor, A.S. (1996) *HoNOS: Health of the Nation Outcome Scales: Report on Research and Development July 1993-December 1995*. London: Royal College of Psychiatrists.
- Wing, J.K., Lelliott, P., & Beevor, A.S. (2000). Progress on HoNOS. *British Journal of Psychiatry*, 176, 392-395. <https://doi.org/10.1192/bjp.176.4.392>
- Wing, JK., Beevor, AS., Curtis, RH., Park, SBG., Hadden, S. and Burns, A. (1998). Health of the Nation Outcome Scales (HoNOS). *British Journal of Psychiatry*, 172, pp. 11-18. **(RevSist)**
- Wise, E.A. (2004). Methods for analyzing psychotherapy outcomes: A review of clinical significance, reliable change, and recommendations for future directions. *Journal of Personality Assessment*, 82(1), 50-59. https://doi.org/10.1207/s15327752jpa8201_10
- Wright, T., Stevens, S., & Woulides, TA. (2018). Mothers and their infants co-admitted to a newly developed mother–baby unit: characteristics and outcomes. *Infant Mental Health Journal*, 39(6), pp. 707-717. doi: 10.1002/imhj.21742 **(RevSist)**
- Wyngaerden, F., Nicaise, P., Dubois, V. et al. (2019). Social support network and continuity of care: an ego-network study of psychiatric service users. *Social Psychiatry and Psychiatric Epidemiology*, 54, 725–735. <https://doi.org/10.1007/s00127-019-01660-7> **(RevSist)**

Y

- Youn, S. J., Kraus, D. R., & Castonguay, L. G. (2012). The treatment outcome package: facilitating practice and clinically relevant research. *Psychotherapy (Chicago, Ill.)*, 49(2), 115–122. <https://doi.org/10.1037/a0027932>

ANEXOS

Anexo 1: Escala HoNOS - Health of the Nation Outcome Scale

Escalas de Evolución HoNOS

Uriarte JJ*, Beramendi V*, Medrano J**, Wing JK***, Beevor AS***, Curtis R***

Versión en Lengua Castellana de: Health of the Nation Outcome Scales

(Wing, Beevor, Curtis)

*Hospital Psiquiátrico de Zamudio. Sº Rehabilitación. Osakidetza-Servicio Vasco de Salud

**Hospital Psiquiátrico Las Nieves. Osakidetza-Servicio Vasco de Salud

***Royal College of Psychiatrists. College Research Unit

Resumen de las Instrucciones de Puntuación

1. **Puntuar** cada una de las escalas siguiendo el orden de la 1 a la 12
2. **No** incluir información ya puntuada en un ítem previo excepto en el ítem 10 que es una evaluación global
3. **Puntuar** el problema MÁS SEVERO ocurrido durante el periodo de tiempo evaluado
4. **Todas las escalas** siguen el formato:

0	= sin problema
1	= problema menor que no requiere intervención
2	= problema leve pero claramente presente
3	= problema de moderada gravedad
4	= problema grave o muy grave

Glosario de la Hoja de Puntuación HoNOS

1. Conducta hiperactiva, agresiva, disruptiva o agitada

Incluye tal comportamiento debido a cualquier causa (por ejemplo, drogas, alcohol, demencia, psicosis, depresión, etc.)

No incluye la conducta bizarra, que se puntúa en la escala 6

- 0.** Sin problema de este tipo durante el periodo evaluado

1. Irritabilidad, enfrentamientos, inquietud etc. que no requiere intervención
2. Incluye actitudes agresivas; empujar o importunar a otras personas; amenazas o agresiones verbales; daños menores a bienes y enseres (por ejemplo, romper vasos, cristales); hiperactividad marcada o agitación
3. Agresividad física hacia otras personas o animales (insuficiente para puntuar 4); actitud amenazante; hiperactividad más grave o destrucción de bienes o enseres
4. Al menos un ataque físico serio a otras personas o animales; destrucción de propiedades (por ejemplo, provocando incendios); conducta intimidatoria u obscena

2. Autolesiones no accidentales

*No incluye las autolesiones accidentales (debidas, por ejemplo, a demencia o a incapacidad intelectual severa); **los problemas cognitivos han de ser puntuados en la Escala 4 y el daño resultante en la Escala 5***

No incluye las enfermedades o lesiones producidas como consecuencia directa del consumo de alcohol o drogas, que se puntúan en la Escala 3; (por ejemplo, la cirrosis hepática o las lesiones derivadas de conducir en estado ebrio se puntúan en la escala 5).

0. Ningún problema de este tipo durante el periodo evaluado
1. Pensamientos pasajeros de acabar con todo, pero riesgo pequeño durante el periodo evaluado; no autolesiones
2. Riesgo leve durante el periodo evaluado; incluye autolesiones no peligrosas (por ejemplo, arañarse las muñecas)
3. Riesgo moderado o grave de autoagresión deliberada durante el periodo evaluado; incluye actos preparatorios (por ejemplo, acumular medicación)
4. Tentativa seria de suicidio y/o autolesión grave deliberada durante el periodo evaluado

3. Consumo Problemático de Alcohol o Drogas

No incluye la conducta agresiva/destructiva secundaria al consumo de alcohol o drogas, puntuado en la Escala 1

No incluye a las enfermedades o discapacidades derivadas del consumo de alcohol o drogas, puntuadas en la Escala 5

0. Ningún problema de este tipo durante el periodo evaluado
1. Consumo excesivo, pero dentro de las normas sociales
2. Pérdida de control en el consumo de alcohol o drogas, pero sin dependencia seria

3. Marcada apetencia o dependencia del alcohol o drogas con frecuente pérdida del control; conductas de riesgo durante los periodos de intoxicación
4. Incapacitado por los problemas de alcohol/drogas

4. Problemas Cognitivos

Incluye problemas de memoria, orientación y comprensión asociadas a cualquier trastorno: incapacidad intelectual, demencia, esquizofrenia, etc.

No incluir trastornos temporales (por ejemplo, resacas) secundarios al consumo de alcohol o drogas, puntuados en la Escala 3

0. Ningún problema de este tipo durante el periodo evaluado
1. Problemas menores de memoria o comprensión (por ejemplo, olvido ocasional de nombres)
2. Problemas leves pero claros (por ejemplo, se ha perdido en un lugar familiar o no ha reconocido a una persona conocida); confusión ocasional en relación con decisiones simples
3. Marcada desorientación en el tiempo, el espacio o las personas; desconcierto ante acontecimientos cotidianos; lenguaje algunas veces incoherente; enlentecimiento mental
4. Desorientación severa (por ejemplo, incapacidad para reconocer a familiares); riesgo de accidentes; habla incomprensible; obnubilación o estupor

5. Problemas por enfermedad física o discapacidad

Incluye enfermedades o discapacidades de cualquier causa que limiten o impidan el movimiento, o deterioren la capacidad visual o auditiva, o que de otra manera interfieran con la autonomía personal

Incluye a los efectos secundarios de la medicación; los efectos del consumo de drogas/alcohol; las discapacidades físicas derivadas de accidentes o de autolesiones asociadas a problemas cognitivos, conducción de vehículos bajo los efectos del alcohol, etc.

No incluye los problemas mentales o conductuales puntuados en la Escala 4

0. Ningún problema de salud física durante el periodo evaluado
1. Problema menor de salud durante dicho periodo (por ejemplo, resfriados, caídas de poca importancia, etc.)

2. Problema de salud física que provoca limitación leve de la movilidad y actividad
3. Grado moderado de limitación de la actividad debido a un problema de salud física
4. Incapacidad severa o completa secundaria a un problema de salud física

6. Problemas asociados a la presencia de ideas delirantes y alucinaciones

Incluye alucinaciones e ideas delirantes independientemente del diagnóstico

Incluye la conducta extraña o bizarra asociada a la presencia de alucinaciones o ideas delirantes

No incluye la conducta agresiva, destructiva o hiperactiva atribuible a la presencia de alucinaciones o ideas delirantes, que se puntúa en la Escala 1

0. No hay evidencia de alucinaciones, ideas delirantes durante el periodo evaluado
1. Creencias excéntricas o en alguna medida extrañas no congruentes con el ambiente cultural del paciente
2. Ideas delirantes o alucinaciones (por ejemplo, voces o visiones) presentes, pero originan escaso malestar o conducta bizarra en el paciente, es decir clínicamente presente, pero de forma leve
3. Preocupación marcada en relación con las ideas delirantes o las alucinaciones, que originan mucho malestar y/o se manifiestan en una conducta bizarra obvia, es decir problema de moderada severidad clínica
4. El estado mental y la conducta del paciente están seriamente y negativamente afectados como consecuencia de las ideas delirantes o de las alucinaciones, con una repercusión severa sobre el paciente

7. Problemas en relación con el humor depresivo

No incluye la hiperactividad o la agitación, que se puntúa en la Escala 1

No incluye la ideación o tentativas de suicidio, que se puntúan en la Escala 2

No incluye la presencia de ideas delirantes o alucinaciones, que se puntúan en la Escala 6

0. Ningún problema asociado con humor depresivo durante el periodo evaluado
1. Tristeza; o cambios menores en el estado de ánimo
2. Depresión y malestar leve pero claro (por ejemplo, sentimientos de culpa; pérdida de la autoestima)
3. Depresión con autoinculpación injustificada; preocupación acerca de sentimientos de culpa
4. Depresión severa o muy severa, con ideas de culpa o autoacusación

8. Otros problemas mentales o conductuales

Puntuar únicamente el problema clínico más severo que no haya sido considerado en los ítems 6 y 7 como sigue. Especificar el tipo de problema consignando la letra apropiada: A fóbico; B ansiedad; C obsesivo-compulsivo; D sobrecarga mental y tensión; E disociativo; F somatoforme; G alimentación; H sueño; I sexual; J otros, especificar

0. Sin evidencia de cualquiera de estos problemas a lo largo del periodo evaluado
1. Problemas menores solamente
2. Problema clínicamente presente con intensidad leve (por ejemplo, el paciente tiene un cierto grado de control)
3. Crisis o malestar severo de forma ocasional, con pérdida de control (por ejemplo, se ve obligado a evitar por completo situaciones que generan ansiedad, pedir ayuda a un vecino, etc.) es decir nivel de problema moderadamente severo
4. Problema grave que domina la mayoría de las actividades

9. Problemas con las relaciones

Puntuar el problema más severo del paciente asociado a la carencia activa o pasiva de relaciones sociales, y/o a relaciones carentes de apoyo, destructivas o autolesivas

0. Ningún problema significativo durante el periodo evaluado
1. Problema menor sin entidad clínica
2. Problema claro para crear o mantener relaciones de apoyo: el paciente se queja de ello y/o tales problemas son evidentes para los demás
3. Problema severo y persistente debido a carencia pasiva o activa de relaciones sociales y/o que las relaciones existentes aportan un escaso o nulo apoyo o consuelo
4. Aislamiento social grave y penoso debido a incapacidad para comunicarse socialmente y/o a la pérdida de las relaciones sociales

10. Problemas con las actividades de la vida cotidiana

Puntuar el nivel medio de funcionamiento en las actividades de la vida diaria (AVD): por ejemplo problemas con actividades básicas de cuidados personales como la alimentación, la limpieza personal, la vestimenta, el uso del lavabo; también habilidades complejas como la administración del dinero, la organización de donde vivir, la ocupación y el tiempo libre, la movilidad y el uso del transporte, las compras, el desarrollo personal, etc.

Incluye: cualquier falta de motivación para manejarse o valerse por sí mismo, puesto que esto contribuye a un menor nivel de funcionamiento

No incluye la falta de oportunidades para ejercer las habilidades y capacidades intactas, que se registra en las Escalas 11 y 12

0. Sin problemas durante el periodo evaluado; buena capacidad de funcionamiento en todas las áreas
1. Problemas menores solamente (por ejemplo, desordenado, desorganizado)
2. Cuidados personales adecuados, pero fallo importante en la realización de una o más habilidades complejas (ver arriba)
3. Problemas severos en una o más áreas de los cuidados personales (alimentación, limpieza personal, vestimenta, uso del lavabo) así como discapacidad severa para ejecutar varias habilidades complejas
4. Discapacidad grave o incapacidad en todas o casi todas las áreas de cuidados personales y habilidades complejas

11. Problemas con las condiciones de vida

Puntuar la severidad media de los problemas en relación con la calidad de las condiciones de vida y con la rutina doméstica cotidiana

¿Están cubiertas las necesidades básicas (calefacción, luz, higiene)? Si es así, ¿se dispone de ayuda para afrontar las discapacidades y de variedad de oportunidades para utilizar habilidades y desarrollar otras nuevas?

No puntuar el nivel de discapacidad funcional en sí mismo, puntuado en la Escala 10

Nota: puntuar el alojamiento habitual del paciente. Si la evaluación se lleva a cabo en una unidad de internamiento de corta estancia, puntuar su alojamiento habitual.

0. El alojamiento y las condiciones de vida son aceptables; útil para mantener cualquier discapacidad registrada en la Escala 10 en el más bajo nivel posible, y para ayudar al paciente a valerse por sí mismo
1. El alojamiento es razonablemente aceptable, aunque existen problemas menores o transitorios (por ejemplo, la situación no es la ideal, no es la opción preferida, la comida no es de su agrado, etc.)
2. Problema significativo con uno o más aspectos del alojamiento y/o régimen (por ejemplo, elección restringida; el personal o las personas que viven con el paciente tienen conocimientos insuficientes acerca de cómo limitar la discapacidad o de cómo ayudar a utilizar o a desarrollar habilidades nuevas o intactas)
3. Múltiples problemas preocupantes en relación con el alojamiento (por ejemplo, carencia de algunas necesidades básicas); los recursos del entorno doméstico para mejorar la independencia del paciente son mínimos o inexistentes

4. El alojamiento es inaceptable (por ejemplo, carece de las necesidades básicas, el paciente está en riesgo de desahucio o de quedarse en la calle, o las condiciones de vida son por lo demás intolerables), agravando los problemas del paciente

12. Problemas en relación con la ocupación y las actividades

Puntuar el nivel global de problemas en relación con la calidad del entorno cotidiano. ¿Existe ayuda disponible para afrontar las discapacidades, y oportunidades para mantener o mejorar capacidades ocupacionales y recreativas y actividades? Considerar factores tales como el estigma, falta de personal cualificado, acceso a recursos de apoyo (por ejemplo, personal y equipamiento de centros de día, talleres, clubes sociales, etc.)

No puntuar el nivel de incapacidad funcional en sí mismo, puntuado en la Escala 10

Nota: Puntuar la situación habitual del paciente. Si el paciente está internado en una unidad de corta estancia, puntuar las actividades durante el periodo previo al ingreso.

0. El entorno cotidiano del paciente es aceptable: útil para mantener cualquier discapacidad puntuada en la Escala 10 en el menor nivel posible, y para ayudar al paciente a valerse por sí mismo.
1. Problema menor o transitorio (por ejemplo, retraso en cobrar el paro): recursos razonables disponibles, pero no siempre en el momento deseado, etc.
2. Surtido limitado de actividades; carencia de tolerancia razonable (por ejemplo, denegar injustificadamente la entrada en bibliotecas u otros lugares públicos, etc.); en desventaja por la falta de domicilio estable; apoyo insuficiente por parte de profesionales o cuidadores; ayuda durante el día disponible, pero durante tiempo muy limitado
3. Deficiencia marcada de servicios cualificados disponibles que ayuden a minimizar el nivel de discapacidad existente; ausencia de oportunidades para utilizar habilidades intactas o para adquirir nuevas; atención no cualificada de difícil acceso
4. La falta de oportunidades para realizar actividades durante el día contribuye a empeorar los problemas del paciente

*Anexo 2: Hoja de Puntuación HoNOS***HoNOS - Hoja De Puntuación****Nombre:****Fecha:**

- | | | |
|--|-----------|----------------------|
| 1. Conducta hiperactiva, agresiva, disruptiva o agitada | 0 1 2 3 4 | <input type="text"/> |
| 2. Autolesiones no accidentales | 0 1 2 3 4 | <input type="text"/> |
| 3. Consumo Problemático de Alcohol o Drogas | 0 1 2 3 4 | <input type="text"/> |
| 4. Problemas cognitivos | 0 1 2 3 4 | <input type="text"/> |
| 5. Problemas por enfermedad física o discapacidad | 0 1 2 3 4 | <input type="text"/> |
| 6. Problemas asociados a la presencia de ideas delirantes y alucinaciones | 0 1 2 3 4 | <input type="text"/> |
| 7. Problemas en relación con el humor depresivo | 0 1 2 3 4 | <input type="text"/> |
| 8. Otros problemas mentales o conductuales
Especificar el tipo de trastorno: A, B, C, D, E, F, G, H, I, J | 0 1 2 3 4 | <input type="text"/> |
| 9. Problemas con las relaciones | 0 1 2 3 4 | <input type="text"/> |
| 10. Problemas en relación con las actividades de la vida cotidiana | 0 1 2 3 4 | <input type="text"/> |
| 11. Problemas con las condiciones de vida | 0 1 2 3 4 | <input type="text"/> |
| 12. Problemas en relación con la ocupación y las actividades | 0 1 2 3 4 | <input type="text"/> |

Puntuación Total:**0 - 48**

Tabla 1. Estructura de la Escala HoNOS

Estructura de la Escala HoNOS		
Título de la sección y nombre del ítem	Rango de puntuación de cada ítem	Rango de puntuación de la sección
A. Problemas Conductuales		
1. Agresividad	0-4	0-12
2. Autoagresiones	0-4	
3. Uso de sustancias	0-4	
B. Deterioro		
4. Disfunción Cognitiva	0-4	0-8
5. Discapacidad física	0-4	
C. Problemas Clínicos		
6. Depresión	0-4	0-12
7. Alucinaciones e I Delirantes	0-4	
8. Otros Síntomas	0-4	
D. Problemas Sociales		
9. Relaciones Sociales	0-4	0-16
10. Funcionamiento General	0-4	
11. Problemas de Residencia	0-4	
12. Problemas Ocupacionales	0-4	
Puntuación Total (1-12)	0-48	

12

Anexo 3: Clasificación Artículos Revisión Sistemática

Autor	Año	País	Localización	N	Población	Estudio	Total				Datos dimensiones
Reino Unido (55 estudios)							α	Media	DE	EEM 95%	
Wing, JK., Beevor, AS., Curtis, RH., Park, SBG., Hadden, S. and Burns, A.	1998	UK (1)	Nottingham/ Manchester	641	SMM	Psicométrico	---	9,98	---	---	
Orrell, M., Yard, P., Handysides, J. and Schapira, R.	1999	UK	Hertfordshire	100	SMM	Psicométrico/ Clínico	0,65	8,54	4,5	0,882	
Bebbington, P., Brugha, T., Hill, T., Marsden, L., and Window S.	1999	UK	London	115	SMM	Psicométrico	---	---	---	---	
Chaplin, R., and Perkins, R.	1999	UK	London	248	SMM	Psicométrico	---	---	---	---	
Amin, S., Singh, S.P., Croudace, T., Jones, P., Medley, I., and Harrison, G.	1999	UK	Nottingham	151	SMM	Psicométrico/ Clínico	---	7,05	5,3	0,845	
Slade, M., Beck, A., Bindman, J., Thornicroft, G., and Wright, S.	1999	UK	---	259	SMM	Psicométrico	---	---	---	---	
Sharma, V.K., Wilkinson, G., and Fear, S.	1999	UK	---	156	SMM	Clínico	---	7,49	---	---	
McClelland, R., Trimble, P., Fox, ML., Stevenson MR., and Bell, B.	2000	UK	Belfast	934	SMM	Psicométrico	---	12,67	---	---	
Kelly, C., Sharkey, V., Morrison, G., Allardyce, J., and McCready, R.G.	2000	UK	Nitchsdale	138	SMG	Clínico	---	---	---	---	
Audin, K., Margison, F.R., Mellor-Clark, J., Barkham, M.	2001	UK	---	1688	SMM	Psicométrico/ Clínico	---	8,93	5,2	0,248	D1- M=1,57, DT=,92; D2-M=0,7, DT=1,2; D3-M=3,84, DT=2,1; D4-M=2,94, DT=2,3
Milne, D., Reichelt, K., and Wood, E.	2001	UK	---	30	Raters	Clínico	---	---	---	---	

Slade, M., Cahill, S., Kelsey, W., Powell, R., and Strathdee, G.	2002	UK	London	274	SMM	Psicométrico/ Clínico	----	9,8	5,1	0,604	
Rees, A., Richards, A., and Shapiro, DA.	2004	UK	Leeds	129	SMM	Psicométrico/ Clínico	---	8,6	5,04	---	
Hunter, R., Mclean, J., Peck, D., Pullen, I., Greenfield, A., McArthur, W., and Quinn, A., Eaglesham, J., Hagen, S., and Norrie, J.	2004	UK	---	695	SMM	Psicométrico/ Clínico	---	11,3	5,7	0,424	
Salvi, G., Leese, M., and Slade, M.	2005	UK	---	264	SMM	Psicométrico/ Clínico	---	9,7	5,1	0,615	
Ecob, R., Croudace, T.J., White, IR., Evans, JE., Harrison, GL., Sharp, D., and Jones, PB.	2006	UK	Bristol	384	SMM	Psicométrico	---	---	---	---	
Hansen, L., and Kingdon, D.	2006	UK	---	40	SMG	Psicométrico	---	---	---	---	
Sugarman, P., Walker, L., and Dickens, G.	2009	UK	Northampton	---	---	Clínico No versión	---	---	---	---	
López-Morinigo, JD., Fernandes, AC., Chang, CK., Hayes, RD., Broadbent, M., Stewart, R., David, AS., and Dutta, R.	2014	UK	London	308	SMG	Clínico	---	9,55	---	---	
Williams, B., Speak, B., Hay P., and Muncer, S.J.	2014	UK	Middlesbrough	2706	SMM	Psicométrico	0,66	---	---	---	
Speak, B., and Muncer, S.	2015	UK	---	80161	SMM	Psicométrico	0,86	---	---	---	
Tulloch, A. D., David, A. S. and Thornicroft, G.	2015	UK	London	7891	SMM	Clínico	---	---	---	---	
Muncer, S., Bass, M., and Dawkin, M.	2016	UK	---	989	SMM	Psicométrico	---	---	---	---	
Speak, BL., and Muncer, SJ.	2016	UK	---	80161	SMM	Psicométrico	---	---	---	---	
Twomey, C., Prina, AM., Baldwin, DS, Das-Munshi, J., Kingdon, D., Koeser, L., Prince, MJ., Stewart, R., Tulloch, AD., Cieza, A.	2016	UK	---	1343	SMM	Clínico	---	11,2	5,39	0,288	
Smith, W., Vitoratou, S., McCrone, P., and Patel, A.	2017	UK	---	12910	SMG	Psicométrico	0,52	---	---	---	
Valerie Moran, V., and Jacobs, R.	2017	UK	---	833	SMM	Clínico	---	11	6	0,407	
Macdonald, A., Adamis, D., Craig, T., and Murray, R.	2018	UK	London	1323	SMG	Clínico	---	10,01	5,26	0,283	

Maddison, P., Golay, P., and Muncer, S.J.	2018	UK	---	2325	SMM	Clínico	---	---	---	---	
Laugharne, R., Eaves, S., Mascas, A., Psatha, K., Dinnis, G., Trower, J., and Shankar, R.	2018	UK	---	1271	SMM	Clínico	---	---	---	---	
Lally, J., Watkins, R., Nash, S., Shetty, H., Gardner- Sood, P., Smith, S., Murray, R.M., and Gaughran, F.	2018	UK	---	293	SMM	Clínico	---	7,16	4,7	0,538	
Stephenson, L.A., Macdonald, A.J.D., Seneviratne, G., Waites, F., and Pawlby, S.	2018	UK	London	152	SMM	Clínico	---	14,07	5,3	0,843	
Taylor, C. L., Stewart, R. J. and Howard, L. M.	2018	UK	London	396	SMM	Clínico	---	---	---	---	
Tucker, S., Hughes, J., Jolley, D., Buck, D., Hargreaves, C. and Challis, D.	2018	UK	London	300	SMM	Clínico	---	---	---	---	
Stubbs, B., Mueller, C., Gaughran, F., Lally, J., Vancampfort, D., Lamb, S. E., Koyanagi, A., Sharma, S., Stewart, R. and Perera, G.	2018	UK	South London	11567	SMM	Clínico	---	11,38	9,47	---	
Hodgson, R.E.	2019	UK	Staffordshire	114	SMM	Clínico	---	15,17	7,31	1,34	
Lei, H., Barnicot, K., Maynard, E., Etherington, A., Zalewska, K., Quirk, A., Sanatinia, R., Cooper, S. J., Crawford, M. J.	2019	UK	---	8760	SMM	Clínico	---	---	---	---	
Gale, T.M. and Boland, B.	2019	UK	---	22164	SMM	Clínico	---	11,36	---	---	
Richardson, T., Dasyam, B., Courtney, H., White, L., Tedbury, J., Butt, J. and Newman-Taylor, K.	2019	UK	---	103	SMM	Psicométrico	0,75	12,9	---	---	
Kugathasan, P., Wu, H., Gaughran, F., Nielsen, R. E., Pritchard, M., Dobson, R., Stewart, R. and Stubbs, B.	2019	UK	South London	9775	SMM	Clínico	---	---	---	---	
Beck, K., McCutcheon, R., Stephenson, L., Schilderman, M., Patel, N., Ramsay, R. and Howes, O.	2019	UK	London	202	SMM	Clínico	---	---	---	---	
Cai, W., Mueller, C., Shetty, H., et al.	2020	UK	London	1292	SMM	Clínico	---	---	---	---	
Dean, N. J., Arnaoutoglou, N. and Underwood, B. R.	2020	UK	Cambridgeshire	6813	SMM	Clínico	---	11,07	---	---	
Crook, H., Duffy, A., Nair, B., & Waters, R.	2022	UK	Easy Kent	40	SMM	Clínico	---	---	---	---	

Chaturvedi, J., Sabbah, W., Gallagher, J. E., Turner, J., Curl, C. and Stewart, R.	2020	UK	London	18999	SMM	Clínico	---	---	---	---	
Hong, J. S. W., Sheriff, R., Smith, K., et al.	2021	UK	Oxford	998	SMM	Clínico	---	---	---	---	
O'Connor, J. and Katyal, R.	2021	UK	Nordfolk	---	SMM	Clínico/ Póster	---	---	---	---	
Taylor, K. M., Orucu, E., Nandha, S. and Cella, M.	2023	UK	London	327	SMM	Clínico	---	11,19	6,39		
Jones, C.I., Bremner, S., Hooper, R. et al.	2021	UK	---	950	SMM	Psicométrico	---	---	---	---	
Mansour, H., Mueller, C., Davis, K. A. S., Burton, A., Shetty, H., et al.	2020	UK	London	13786	SMG	Clínico	---	---	---	---	
Cranshaw, J., Seneviratne, G., Rao, R., Ogunmuyiwa, J., McMillin, R., Ntephe, C., and Dain, V.	2021	UK	London	114	SMM	Clínico/ Póster	---	12,83	---	---	
Penington, E., Lennox, B., Geulayov, G., Hawton, K. and Tsiachristas, A.	2022	UK	Oxford	51880	SMM	Clínico	---	11,6	6,3	---	
Mascio, A., Stewart, R., Botelle, R., Williams, M., Mirza, L., Patel, R., Pollak, T., Dobson, R. and Roberts, A.	2021	UK	---	24614	SMM	Clínico	---	---	---	---	
Gale, T. M. and Boland, B.	2021	UK	---	726	SMM	Clínico	---	13,53	7,10	---	
Monk-Cunliffe, J., Kadra-Scalzo, G., Finamore, C., Dale, O., Khondoker, M., Barrett, B., ... Moran, P.	2023	UK	---	6912	SMM	Clínico	---	---	---	---	
Nueva Zelanda (8 estudios)											
Eagar, K., Trauer, T., and Mellsop, G.	2005	NZ (2)	---	12576	SMM	Psicométrico	0,80	---	---	---	
Mellsop, G., Dutu, G., and El-Badri, S.	2007	NZ	---	---	SMG	Clínico	---	---	---	---	
Stewart, M.	2009	NZ	Auckland	25	SMM	Psicométrico/ Clínico	---	8	7,1	2,783	D1- M=1.2, DT=2.3; D2-M=1, DT=1.4; D3-M=2.5, DT=2.7; D4-M=3.3, DT=3
Beaglehole, B., Frampton, CM., Boden, JM., Mulder, RT., and Bell, CJ.	2017	NZ	Canterbury	---	SMM	Clínico	---	---	---	---	
Jury, A., Lai, J., Tuason, C., Koning, A., Smith, M., Boyd, L., Swanson, C., Fergusson, D., and Guar, A.	2018	NZ	---	842	SMM	Psicométrico	---	---	---	---	

Wright, T., Stevens, S., and Wouldes, TA.	2018	NZ	---	45	SMM	Psicométrico /Clínico	---	13,87	5,18	1,513	
Lai, J., Jury, A., Long, J., Fergusson, D., Smith, M., Baxendine, S., and Gruar, A.	2018	NZ	---	10727	SMM	Clínico	---	14,529	---	---	
Barbalat, G., Van den Bergh, D. and Kossakowski, J.J.	2019	NZ	Auckland	137	SMM	Clínico	---	---	---	---	
Australia (69 estudios)											
Goldney, RD., Fisher, LJ., and Walmsley, SH.	1998	Australia (3)	---	819	SMM	Psicométrico/ Clínico	---	12,6	6,4	0,438	
Trauer, T.	1999	Australia	---	2137	SMM	Psicométrico	0,75	---	---	---	
Trauer, T., Callaly, T., Hantz, P., Little, J., Shields, R., and Smith, J.	1999	Australia	Victoria	2137	SMM	Psicométrico	---	---	---	---	
Issakidis, C., and Teesson, M.	1999	Australia	---	123	SMM	Psicométrico/ Clínico	---	10,1	7	1,237	
Trauer, T., Callaly, T., and Hantz, P.	1999	Australia	Geelong	477	SMM	Clínico	---	15,09	6,21	0,557	
Gallagher, J., and Teesson, M.	2000	Australia	---	260	SMM	Psicométrico/ Clínico	0,76	9,88	5,8	0,705	
Brooks, R.	2000	Australia	New South Wales	318	SMM	Psicométrico	---	---	---	---	
Preston, NJ.	2000	Australia	Perth	1957	SMM	Psicométrico	---	---	---	---	
Kisely, S., Preston, N., and Rooney, M.	2000	Australia	---	2691	SMM	Clínico	---	---	---	---	
Hugo, M.	2000	Australia	---	678	SMM	Clínico	---	14,86	5,74	0,432	
Stedman, T., Yellowlees, P., Drake, S., Chant, D., Clarke, R., and Chapple, B.	2000	Australia	---	173	SMM	Clínico	---	19,06	0,56	0,083	
Rock, D. and Preston, N.	2001	Australia	Perth	142	---	Psicométrico	---	---	---	---	
Page, A.C., R. Hooke, G. R., M. Rutherford, E. M.	2001	Australia	---	747	SMM	Psicométrico/ Clínico	0,59	12,3	5	0,359	
Kennedy, C., and Yellowlees, P.	2002	Australia	Queensland	124	SMM	Clínico	---	11,1	5,8	1,021	
Hooke, G.R., and Page, A.C.	2002	Australia	---	348	SMM	Clínico	---	---	---	---	
Hugo, M., Smout, M., and Bannister, J.	2002	Australia	---	514	SMM	Clínico	---	---	---	---	

Nash, L., Gorrell, J., Cornish, A., Rosen, A., Miller, V., and Tennant.	2004	Australia	Sydney	57	SMG	Clínico	---	13,9	5	1,298	
Trauer, T.	2004	Australia	Melbourne	3351	SMM	Clínico	---	---	---	---	
Trauer, T.	2004	Australia	Melbourne	1582	SMM	Clínico	---	15,97	6,53	0,322	
Chopra, P.K., Couper, J.W., Herrman, H.	2004	Australia	---	20	SMM	Psicométrico	---	12,5	---	---	
Waghorn, G., Chant, D., and King, R.	2005	Australia	---	92	SMM	Psicométrico	---	---	---	---	
Newnham, E.A., Harwood, K.E., and Page, A.C.	2006	Australia	---	1830	SMM	Clínico	---	---	---	---	
Bradley, G.M., Couchman, G.M., Perlesz, A., Nguyen, A.T., Singh, B., and Riess.	2006	Australia	Melbourne	25	SMG	Clínico	---	11	5,82	2,281	
Steel, Z., McDonald, R., Silove, D., Bauman, A., Sandford, P., Herron, J., and Minas, I.H.	2006	Australia	---	146	SMM	Clínico	---	15	5,2	0,843	
Wade, D., Harrigan, S., Harris, M.G., Edwards, J., and McGorry, P.D.	2006	Australia	---	104	SMM	Clínico	---	---	---	---	
Green, H.J., Frank, R.M., Butwell, A., and Beck, O.J.	2007	Australia	---	42	SMM	Clínico	---	13,5	4,2	1,27	
McNaught, M., Caputi, P., Oades, L.G., and Deane, F.P.	2007	Australia	---	168	SMM	Psicométrico	---	---	---	---	
Miller, K., Walterfang, M., Randhawa, S., Scholes, A., Mocellin, R., and Velakoulis, D.	2007	Australia	---	76	SMM	Psicométrico/ Clínico	---	17,18	6,13	1,378	
Berk, M., Ng, F., Dodd, S., Callaly, T., Campbell, S., Bernardo, M., and Trauer, T.	2007	Australia	Victoria	510	SMM	Psicométrico	---	9,21	---	---	
Murugesan, G., Amey, C.G., Deane, F.P., Jeffrey, R., Kelly, B., and Stain, H.	2007	Australia	---	84	SMM	Clínico	---	15,82	6,29	1,345	
Prabhu, R., and Oakley Browne, M.	2008	Australia	Victoria	122	SMM	Psicométrico	---	---	---	---	
Happell, B., Sundram, S., Wortans, J., Johnstone, H., Ryan, R., and Lakshmana, R.	2009	Australia	---	51	SMM	Clínico	---	12,2	4,8	1,317	
Newnham, E.A., Harwood, K.E., and Page, A.C.	2009	Australia	---	6475	SMM	Psicométrico/ Clínico	0,45	13,17	4,98	0,121	D1 alpha=0,43; D2 alpha=0,37; D3 alpha=0,55; D4 alpha=0,45
Burgess, P., Trauer, T., Coombs, T., McKay, R., and Pirkis, J.	2009	Australia	---	94	---	Psicométrico	---	---	---	---	

Burgess, P., Pirkis, J., and Coombs, T.	2009	Australia	---	>20000	SMM	Clínico	---	---	---	---	
Kelly, P.J., Deane, F.P.	2009	Australia	---	129	SMM	Psicométrico	0,72	---	---	---	
Duke, B.	2010	Australia	Queensland	331	---	Clínico	---	10,1	6,78	0,73	
Callaly, T., Trauer, T., Hyland, M., Coombs, T. and Berk, M.	2011	Australia	---	368	SMM	Clínico	---	14,9	---	---	
Gilbert, M.M., Chamberlain, J.A., White, C.R., Mayers, P.W., Pawsey, B., Liew, D., Musgrave, M., Crawford, K., and Castle, D.J.	2012	Australia	---	44	SMM	Clínico	---	7,25	4,7	1,389	
Trauer, T., and Nagel, T.	2012	Australia	Melbourne	271	SMM	Clínico	---	11,8	---	---	
Brann, P., Alexander, M., and Coombs, T.	2012	Australia	---	49	---	Psicométrico/ Clínico	---	16,5	4,9	1,372	
Gonda, T., Deane, F.P., and Murugesan, G.	2012	Australia		329	SMM	Clínico	---	13,49	5,96	0,644	
Siskind, D., Harris, M., Kisely, S., Brogan, J., Pirkis, J., Crompton, D., and Whiteford, H.	2013	Australia	Queensland	---	SMM	Clínico	---	---	---	---	
Coombs, T.	2014	Australia	---	94145	SMM	Clínico	---	14,1	6,7	0,043	
Kisely, S., Xiao, J., Crowe, E., Paydar, A., Jian, L.	2014	Australia	---	1433	SMM	Clínico	---	10,6	9,7	0,502	
Siskind, D., Harris, M., Diminic, S., Carstensen, G., Robinson, G., and Whiteford, H.	2014	Australia	---	1757	SMM	Clínico	---	17,58	0,25	0,012	
McKay, R., and Coombs, T.	2012	Australia	---	---	SMM	Clínico	---	---	---	---	
Parker, S., Dark, F., Newman, E., Korman, N., Meurk, C., Siskind, D., and Harris, M.	2016	Australia	---	S.D.	SMM	Clínico	---	---	---	---	
Luo, W., Harvey, R., Tran, T., et al.	2016	Australia	Victoria	111	SMM	Psicométrico	0,70	12,16	---	---	
Parker, S., Dark, F., Newman, E., Hanley, D., McKinlay, W. and Meurk, C.	2017	Australia	---	24	SMM	Clínico	---	12	6,4	2,561	
Penno, S. J., Hamilton, B. and Petrakis, M.	2017	Australia	---	270	SMM	Clínico	---	11,86	6,12		
Gore-Jones, V., and Dark, F.	2018	Australia	---	249	SMM	Clínico	---	14,32	7,54	0,937	
O'Connell, J., Sunwoo, M., McGorry, P. and O'Donoghue, B.	2018	Australia	---	544	SMM	Clínico	---	---	---	---	

Maxwell, A., Tsoutsoulis, K., Tarur Padinjareveettil, A. M., Zivkovic, F. and Rogers, J. M.	2018	Australia	Sidney	173	SMG	Psicométrico/ Clínico	---	6,46	5,1	---	
Bardell-Williams, M., Eaton, S., Downey, L., Bowtell, M., Thien, K., Ratheesh, A., Killackey, E., McGorry, P. and O'Donoghue, B.	2019	Australia	---	544	SMM	Clínico	---	---	---	---	
Segal, S. P., Rimes, L., and Hayes, S. L.	2019	Australia	Victoria	27585	SMM	Clínico	---	---	---	---	
Parker, S., Siskind, D., Hermens, D. F., Dark, F., McKeon, G., Korman, N., Arnautovska, U., Harris, M. and Whiteford, H.	2019	Australia	Brisbane	144	SMM	Clínico	---	10,31	5,99	---	D1- M=1,13, DT=1,61; D2M=1,72,DT=1,52; D3-M=3,56, DT=2,41; D4-M=3,9, DT=2,97
Sunwoo, M., O'Connell, J., Brown, E., Lin, A., Wood, S. J., McGorry, P. and O'Donoghue, B.	2019	Australia	Melbourne	544	SMM	Clínico	---	---	---	---	
Seow, L. L. Y., Page, A. C. and Hooke, G. R.	2020	Australia	---	102	SMM	Psicométrico/ Clínico	0.65	9,38	2,31	---	
Parker, S., Arnautovska, U., Siskind, D., Dark, F., McKeon, G., Korman, N., and Harris, M.	2020	Australia	Queensland	501	SMM	Psicométrico/ Clínico	---	18,44	8,17	---	
Irvine, A., Rawlinson, C., Bor, W., & Hoehn, E.	2021	Australia	Queensland	84	SMM	Clínico	---	8,35	3,59	---	
Whitehead, R., Hopkins, L., Hughes, E., Kehoe, M. and Pedwell, G.	2020	Australia	Melbourne	67	SMM	Clínico	---	19,30	7,48	---	D1- M=4,74, DT=3,24; D2-M=3,22, DT=1,47; D3-M=2,00, DT=2,15; D4-M=9,35, DT=3,90
Soni, N., Roberts, S. and Branjerdporn, G.	2022	Australia	Queensland	82	SMM	Clínico	---	11,10	6,02	---	
Branjerdporn, G., Robertson, J. and Dymond, A.	2022	Australia	Queensland	130	SMM	Clínico	---	15,26	8,78	---	D1- M=4,67, DT=2,97; D2-M=0,77, DT=1,15; D3-M=5,00, DT=2,49; D4-M=4,82, DT=4,70
Mukherjee, H. and Sazhin, V.	2022	Australia	Sydney	62	SMM	Clínico	---	---	---	---	
Sekharan, L., Jagadheesan, K., Das, P., Lakra, V., West, K. and Baldwin, M.	2021	Australia	Melbourne	67	SMM	Clínico	---	11,80	6,50	---	
Parker, S., Arnautovska, U., McKeon, G. and Kisely, S.	2021	Australia	Queensland	311	SMM	Clínico	---	12,99	9,14	---	
Charles, R., Bhandarkar, R. and Fulgoni, C.	2023	Australia	Monash	24	SMM	Clínico	---	---	---	---	

Dhar, A., Edwards, L. and Waerea, M.	2021	Australia	Monash	212	SMM	Clínico	---	15,90	---	---	
España (14 estudios) / Uruguay (1 estudio)											
Gisbert, C., and Cid Colom, J.	2010	España (4)	Cataluña	831	SMM	Revisión/ Clínico	---	17,7	5,76	0,392	
Frades, B., Ruipérez, M.A., and Moro, M.	2011	España	Valencia	214	SMG	Clínico	---	20,1	5,75	0,77	
Osta Samanes, R., García Tejedor, A., and Muro, C.	2012	España	Zaragoza	16	SMM	Clínico	---	12,8	---	---	
Vargas, M.L., Quintanilla, P., Lozano, A., Sendra-Gutiérrez, J.M.	2013	España	Castilla y León	74	SMG	Psicométrico	---	---	---	---	
Dogmanas, D.	2014	Uruguay (5) /España	Salamanca	402	SMM	Psicométrico/ Clínico	0,69	7,6	4,2	0,411	
Escalada-Hernández, P., Muñoz-Hermoso, P., González-Fraile, E., Santos, B., González-Vargas, J.A., Feria-Raposo, I., and Girón-García, J.L., and García-Manso, M.	2014	España	Navarra	690	SMG	Clínico	---	13,24	5,97	0,445	
Garay Arostegui, M., Pousa Rodríguez, V., Pérez Cabeza, L.	2014	España	Bizkaia	59	SMG	Clínico	---	9,05	3,56	0,908	
Escalada-Hernández, P., Lazkanotegui-Matxiarena, U., Aolondriz-Urrutia, M., Alameda-Blanco, M., Flores del Redal-López, C., Dendarrieta-Bardot, M.	2015	España	Navarra	108	SMG	Clínico	---	16,81	6,06	1,143	
Escalada-Hernández, P., and Marín-Fernández, B.	2015	España	Navarra	694	SMM	Clínico	---	13,4	4,6	0,342	
Ribas, S., Infante, P., and Fernández-González, D.	2016	España	Barcelona	63	SMG	Clínico / Póster	---	---	---	---	
Lascorz, D., Lopez, V., Pinedo, C., Trujols, J., Vegué, J., and Perez, V.	2016	España	Barcelona	139	SMM	Psicométrico	---	---	---	---	
Vázquez-Bourgon, J., Gómez Ruiz, E., Hoyuela Zátón, F., Salvador Carulla, L., Ayesa Arriola, R., Tordesillas Gutiérrez, D. and Crespo Facorro, B.	2019	España	Santander	331	SMM	Clínico	---	10,13	3,95	0,426	

Martín-Pérez, C., Herrero-Martin, J. J., Pérez-López, G., Pedrosa-García, R., Ibañez-Casas, I., Torres-González, F. and Moreno-Küstner, B.	2019	España	Granada	260	SMM	Clínico	---	9,1	---		
Martin-Iñigo, L., Ortiz, S., Urbano, D. et al.	2022	España	Barcelona	81	SMM	Clínico	---	20,13	4,90	---	
Uriarte, J. J., Iglesias, N., Penas, P., Moreno-Calvete, M. C., Álvarez-González, A. and Iraurgi, I.	2024	España	Bilbao	281	SMG	Psicométrico/ Clínico	0,7	11,30	5,98		
Italia (14 estudios)											
D'Avanzo, B., Battino, R.N., Gallus, S., and Barbato, A.	2004	Italia (6)	Milán	1792	SMM	Clínico	---	---	---	---	
Parabiaghi, A., Barbato, A., D'Avanzo, B., Erlicher, A., and Lora, A.	2005	Italia	Lombardy	9817	SMM	Psicométrico/ Clínico	0.73	8,7	5,6	0,111	
Barbato, A., Agnetti, G., D'Avanzo, B., Frova, M., Guerrini, A., Tettamanti, M.	2007	Italia	Milán	103	SMM	Clínico	---	19,5	6,6	1,275	
Cocchi, A., Meneghelli, A., and Preti, A.	2008	Italia	Milán	143	SMM	Clínico	---	14,16	5,87	0,962	
Lovaglio, PG., and Monzani, E.	2011	Italia	Milán	2162	SMM	Psicométrico	0.71	9,81	---	---	
Parabiaghi, A., Rapisarda, F., D'Avanzo, B., Erlicher, A., Lora, A., and Barbato, A.	2011	Italia	Milán	2059	SMM	Clínico	---	10,8	6,2	0,268	
Preti, A., Pisano, A., Cascio, MT., Galvan, F., Monzani, E., Meneghelli, A., and Cocchi, A.	2012	Italia	Milán	168	SMG	Psicométrico/ Clínico	0,70	14,2	5,9	0,892	
Lovaglio, PG., and Monzani, E.	2012	Italia	Milán	1633	SMM	Psicométrico/ Clínico	---	10,6	6,2	0,301	
Vita, A., Deste, G., Barlati, S., De Peri, L., Giambra, A., Poli, R., Keefe, RSE., and Sacchetti, E.	2013	Italia	---	86	SMG	Psicométrico	0,78	---	---	---	
Deste, G., Barlati, S., Gregorelli, M., Lisoni, J., Turrina, C., Valsecchi, P. and Vita, A.	2018	Italia	---	75	SMM	Clínico	---	14,67	5,21	1,179	
Elvederi Murri, M., Bertelli, R., Carozza, P., et al.	2020	Italia	Ferrara	86	SMM	Clínico	---	21,90	9,00	---	
Fontecedro, E., Furlan, M., Tossut, D., Pascolo-Fabrizi, E., Balestrieri, M., Salvador-Carulla, L., D'Avanzo, B. and Castelpietra, G.	2020	Italia	---	128	SMM	Clínico	---	13,35	7,70	---	

Elisa, F., Davide, T., Luisa, S. M., Martina, M., Tommaso, B., Luis, S. C., ... Giulio, C.	2023	Italia	---	330	SMM	Clínico	---	11,54	6,40	---	
Pelizza, L., Maestri, D., Leuci, E., Paulillo, G., Ceroni, P., Quattrone, E., and Pellegrini, P.	2022	Italia	Parma	137	SMM	Clínico	---	---	---	---	
Países Bajos (27 estudios)											
Mulder, C.L., Staring, A.B.P., Loos, J., Buwalda, V.J.A., Kuijpers, D., Sytema, S., and Wierdsma A.I.	2004	Países Bajos (7)	---	559	SMM	Psicométrico/ Clínico	0,78	11,4	7	0,580	
Tenneij, N., Didden, R., Veltramp, E., and Koot, HM.	2009	Países Bajos	---	79	SMM	Psicométrico/ Clínico	0,48/ 0,74	10,9/ 12,6	4,5/ 6,6	0,99/ 1,455	
Buwalda, V.J.A., Draisma, S., Smit, J.H., Swinkels, J.A., Van Tilburg, W.	2011	Países Bajos	---	148	SMM	Psicométrico	0,64	---	---	---	
Kortrijk, H.E., Mulder, C.L., Drukker, M., Wiersma, D., and Duivenvoorden, H.J.	2012	Países Bajos	Rotterdam	569	SMM	Clínico	---	18,04	5,24	0,431	
Kortrijk, H. E., Mulder, C. L., van der Gaag, M. and Wiersma, D.	2012	Países Bajos	Rotterdam	278	SMG	Clínico	---	---	---	---	
Fassaert, T., Lauriks, S., van de Weerd, S., Theunissen, J., Kikkert, M., Dekker, J., Buster, M., and de Wit, M.	2014	Países Bajos	Rotterdam	107	SMM	Psicométrico	---	---	---	---	
De Beurs, E., Blankers, M., Delespaul, P., Van Duijn, E., Mulder, N., Nugter, A., Swildens, W., Tiemens, B.G., Theunissen, J., Van Voorst, A.F.A., and Van Weeghel, J.	2018	Países Bajos	---	16671	SMM	Clínico	---	12,4	6,5	0,099	
Stiekema, A.P.M., Looijmans, A., Van Der Meer, L., Bruggeman, R., Schoevers, R.A., Corpeleijn, E., and Jörg, F.,	2018	Países Bajos	---	562	SMM	Clínico	---	13,4	6,2	0,513	
Noordraven, E.L., Wierdsma, A.I., Blanken, P., Bloemendaal, A.F.T., and Mulder, C.L.	2018	Países Bajos	---	169	SMG	Clínico	---	---	---	---	
Deenik, J., Tenback, D.E., Tak, E.C.P.M., Hendriksen, I.J.M., and van Harten, P.N.	2018	Países Bajos	---	112	SMM	Clínico	---	16,6	5,51	1,02	

Jochems, E.C., Duivenvoorden, H.J., van Dam, A. et al.	2018	Países Bajos	---	294	SMM	Psicométrico	0,70	9	---	---	
Albers, W. M. M., Roeg, D. P. K., Nijssen, Y., van Weeghel, J., Bongers, I. M. B.	2018	Países Bajos	---	395	SMM	Psicométrico	0,78	10,81	---	---	
Kortrijk, H.E., Mulder, N.L., Kamperman, A.M. et al.	2018	Países bajos	---	350	SMM	Clínico	---	12,02	6,32	0,662	
Swildens, W. E., Visser, E., Bähler, M., Bruggeman, R., Delespaul, P., van der Gaag, M., de Haan, L., Keet, R., Nijssen, Y., van Os, J., Pijnenborg, G. M., Slooff, C., de Vos, A., van Weeghel, J., Wunderink, L., Mulder, C. L., and Wiersma, D.	2018	Países Bajos	---	840	SMM	Clínico	---	---	---	---	
Kortrijk, H., Schaefer, B., van Weeghel, J., Mulder, C.L., and Kamperman, A.	2019	Países Bajos	---	2660	SMM	Psicométrico/ Clínico	---	12,7	6,3	0,239	
Sanches, S.A., Swildens, W.E., van Busschbach, J.T. et al.	2019	Países Bajos	---	1069	SMM	Clínico	---	7,15	0,16	0,01	
de Vries, B., Pijnenborg, G. H. M., van der Stouwe, E. C. D., Visser, E., de Jong, S., Pharmacotherapy and outcome survey (PHAMOUS)-investigators, et al.	2019	Países Bajos	---	353	SMM	Clínico	---	---	---	---	
van Kranenburg, G.D., van den Brink, R.H.S., Mulder, W.G. et al.	2019	Países Bajos	---	229	SMM	Clínico	---	---	---	---	
Jansen, L., Hunnik, F., Busschbach, J.J.V., Lijmer, J.G.	2019	Países Bajos	---	50	SMM	Psicométrico/ Clínico	0,63	16,2	8	2,217	
Kwakernaak, S., Swildens, W.E., van Wel, T.F. et al.	2019	Países Bajos	Utrecht	287	SMM	Psicométrico	0,75	---	---	---	
Verheesen, S. M. H., ten Doesschate, F., van Schijndel, M. A. et al.	2021	Países Bajos	Arnhem	24	SMM	Clínico	---	20,60	6,90	---	
Liemburg, E., Nienhuis, F. and Veling, W.	2020	Países Bajos	Groningen	164	SMM	Psicométrico	---	---	---	---	

Sportel, B. E., Aardema, H. and Boonstra, N. et al.	2023	Países Bajos	---	259	SMM	Psicométrico/ Clínico	---	8,61	5,62	---	
Bartels-Velthuis, A., Ties, K., Visser, E., Arends, J., Pijnenborg, M., Wunderink, L., Jörg, F., Veling, W., Castelein, S., Knegtering, H. and Bruggeman, R.	2020	Países Bajos	Groningen	156	SMM	Psicométrico	---	---	---	---	
Bruins, J., Pijnenborg, G. H. M., Visser, E. and Castelein, S.	2021	Países Bajos	---	2994	SMM	Clínico	---	---	---	---	
Noorthoorn E.O., Smits, H. J. H., Penterman, E. J. M., Seelen-de Lang, B. L., Nieuwenhuis, J. G., Nijman, H. L. I.	2021	Países Bajos	---	577	SMM	Clínico	---	---	---	---	
Veereschild, H. M., Noorthoorn, E. O., Nijman, H. L. I., Mulder, C. L., Dankers, M., Van der Veen, J. A., Loonen, A. J. M. and Hutschemaekers, G. J. M.	2020	Países Bajos	---	83	SMM	Clínico	---	---	---	---	
Dinamarca (4 estudios)											
Bech, P., Bille, J., Schüte, T., Søndergaard, S., Waart, S., and Wiese, M.	2003	Dinamarca (8)	---	963	SMM	Psicométrico	---	11,36	---	---	D4-M=5,75, DT=3,6
Hansen, JP., Østergaard, B., Nordentoft, M., and Hounsgaard, L.	2012	Dinamarca	---	62	SMG	Clínico	---	---	---	---	
Dinesen Østergaard, S., Bille, J., Søltoft-Jensen, H., Lauge, N., and Bech, P.	2012	Dinamarca	---	357	SMM	Clínico	---	11,01	5,48	0,568	
Nordbrandt, M. S., Sonne, C., Mortensen, E. L. and Carlsson, J.	2020	Dinamarca	Capital Region	318	SMM	Clínico	---	---	---	---	
Noruega (4 estudios)											
Øiesvold, T., Bakkejord, T., abd Sexton, JA.	2011	Noruega (9)	---	118	SMM	Psicométrico/ Clínico	0,67	9,49	0,44	0,079	D1 alpha=0,57; D2 alpha=0,18; D3 alpha=0,27; D4 alpha=0,53
Strømme, M. F., Bartz-Johannesen, C., Kroken, R. A., Mehlum, L. and Johnsen, E.	2022	Noruega	Bergen	426	SMM	Clínico	---	---	---	---	

Lund-Heimark, H., Kjelby, E., Mehlum, L., Gjestad, R., Selbæk, G., Kroken, R. A., Johnsen, E., Oedegaard, K. J. and Mellesdal, L. S.	2020	Noruega	Bergen	917	SMM	Clínico	---	12,06	12,5	---	
Skar-Fröding, R., Clausen, H., Šaltytė Benth, J., Ruud, T., Slade, M. and Heiervang, K. S.	2022	Noruega	---	318	SMM	Clínico	---	---	---	---	
Finlandia (1 estudio)											
Leijala, J., Kampman, O., Suvisaari, J. et al.	2021	Finlandia (10)	---	276	SMG	Clínico	---	7,2	4,5	---	D1- M=0,3, DT=0,7; D2-M=1,8, DT=1,6; D3-M=3,0, DT=2,2; D4-M=1,2, DT=1,3
Alemania (5 estudios)											
Andreas, S., Harfst, T., Dirmaier, J., Kawski, S., Koch, U., & Schulz, H.	2007	Alemania (11)	Hamburgo	1918	SMM	Psicométrico/ Clínico	---	16,56	1,03	0,046	
Andreas, S., Harfst, T., Rabung, S., Mestel, R., Schauenburg, H., Hausberg, M., Kawski, S., Koch, U., and Schulz, H.	2009	Alemania	---	3169	SMM	Psicométrico/ Clínico	0,62	16,99	6,31	0,220	
Andreas, S., Harries-Hedder, K., Schwenk, W., Hausberg, M., Koch, U., and Schulz, H.	2010	Alemania	---	417	SM M	Psicométrico	0,73	---	---	---	
Andreas, S., Rabung, S., Mestel, R., Koch, U., Hausberg, M., and Schulz, H.	2011	Alemania	---	1918 + 55	SMM	Psicométrico	---	---	---	---	
Listunova, L., Bartolovic, M., Kienzle, J., Jaehn, A., Grützner, T. M., Wolf, R. C., Weisbrod, M. and Roesch-Ely, D.	2020	Alemania	---	62	SMM	Clínico	---	---	---	---	
Suiza (18 estudios)											
Theodoridou, A., Jaeger, M., ketteler, D., Kawohl, W., Lauber, C., Hoff, P., and Rössler, W.	2009	Suiza (12)	Zurich	100	SMM	Psicométrico/ Clínico	---	20,6	5,9	1,156	
Kocha, N., Wasserfallenb, JB., Camusa, D., Gebela, S., Bonsacka, C.	2013	Suiza	Lausanne	1737	SM M	Psicométrico/ Clínico	---	15,76	10,4	0,489	

[illegible]

Lauzon, S., Corbière, M., Bonin, JP., Bonsack, C., Lesage, AD., and Ricard, N.	2001	Canadá (13)	Québec	194	SMM	Psicométrico	0,70	---	---	---	
Kisely, S., Campbell, LA., Crossman, D., Gleich, S., and Campbell, J.	2007	Canadá	Nueva Escocia	318	SMM	Psicométrico/ Clínico	---	11,4	5,59	0,614	
Kisely, S., Campbell, LA., Cartwright, J., Cox, M., and Campbell, J.	2010	Canadá	---	4620	SMM	Psicométrico/ Clínico	---	10	5,5	0,159	
Currie, S.R., Liu, P., Adamyk-Simpson, J. et al.	2019	Canadá	Alberta	73	SMM	Psicométrico/ Clínico	---	12,2	5,4		
Irlanda (5 estudios)											
Browne, S., Doran, M., and McGauran, S.	2000	Irlanda (14)	Dublín	100	SMM	Psicométrico/ Clínico	---	11,8	6,3	1,235	
O'Brien, S., McFarland, J., Kealy, B., Püllela, A., Saunders, J., Cullen, W., and Meagher, D.	2012	Irlanda	Limerick	80	SMM	Clínico	---	10,5	4,6	1,008	
Mulligan, O., Helmi, L., Adamis, D., O'Mahony, E., McCarthy, G., and MacDonald, A.	2016	Irlanda	Sligo	280	SMG	Psicométrico/ Póster	---	---	---	---	
Sweeney, J., Adamis, D., Helmi, L. et al.	2019	Irlanda	Sligo/Leitrim	470	SMM	Psicométrico/ Clínico	---	4,43	4,36	0,394	
Crowley, S., McDonagh, S., Carolan, D. and O'Connor, K.	2023	Irlanda	Cork	1041	SMM	Clínico	---	13,60	---	---	
República Checa (4 estudios)											
Pěč, O., Čechová, D., Pěčová, J., Dragomirecká, E., Šelepová, P., Probstová, V., and Bubela, M.	2009	Rep Checa (15)	Praga	40	SMM	Psicométrico / Clínico	0,79	7,85	5,27	1,633	D1- M=1,45, DT=0,89; D2-M=1,3, DT=1,34; D3-M=1,9, DT=1,65; D4-M=3,2, DT=2,46
Pec, O., Bob, P., Pec, J., and Hrubcova., A.	2017	Rep Checa	---	60	SMG	Clínico	---	10,88	---	---	
Páv, M., Vnřuková, M. and Sebal, I.	2022	Rep Checa	Praga	260	SMM	Clínico	---	2,05	2,71	---	D1- M=0,25, DT=0,77; D2-M=0,65, DT=1,18; D3-M=0,63, DT=1,06; D4-M=0,52, DT=1,20

Machova´, K., Juri´ckova´, V., Kasparova´, A., Petrova´, K., Sla´dkova´, B. and Svobodova´, I.	2023	Rep Checa	Kosmonosy	57	SMM	Psicométrico/ Clínico	0,52	11,31	2,97		D1- M=4,3, DT=1,2; D2-M=1,09, DT=1,28; D3-M=2,19, DT=1,13; D4-M=17,42,DT=4,5
USA (1 estudio)											
Muehsam, J.P.	2018	USA (16)	Pennsylvania	793	SMM	Clínico	---	---	---	---	
Bélgica (2 estudios)											
Lorant, V., Grard, A., Van Audenhove, C. et al.	2019	Bélgica (17)	Leuven	1407	SMM	Clínico	---	12	6,8	0,355	
Wyngaerden, F., Nicaise, P., Dubois, V. et al.	2019	Bélgica	Bruselas	380	SMM	Clínico	---	11,6	6,7	0,674	
Hong Kong (1 estudio)											
Hariman, K., Cheng, K. M., Lam, J., Leung, S. K., Lui, S. S. Y.	2020	Hong Kong (18)	---	27637	SMM	Clínico	---	---	---	---	
Grecia (3 estudios)											
Peritogiannis, V., Grammeniati, A., Gogou, A., Gioti, P. and Mavreas, V.	2019	Grecia (19)	---	55	SMM	Clinico	---	---	---	---	
Georgaca, E., Machaira, S., Stamovlasis, D., Peppou, L. E., Papachristou, C., Arvaniti, A., Samakouri, M., Stylianidis, S., Bozikas, V. P., Diakogiannis, I., Fokas, K., Garyfallos, G., Gkolia, I., Karpouza, V., Nimatoudis, I., Patsinakidis, G., Sevrís, D. and Vlachaki, A.	2022	Grecia	Tesalónica	1067	SMM	Clinico	---	---	---	---	
Drakonakis, N., Stylianidis, S., Peppou, L.E. et al.	2021	Grecia	---	544	SMM	Psicométrico/ Clínico	---	16,28	5,96	---	
Indonesia (1 estudio)											
Anjara, S.G., Bonetto, C., Ganguli, P., Setiyawati, D., Mahendradhata, Y., Yoga, B. H., et al.	2019	Indonesia (20)	---	287	SMM	Psicométrico/ Clínico	---	5,45	5,13	---	
		11 idiomas									

Anexo 4: Clasificación Artículos Psicométricos Revisión Sistemática

Autor	Año	País	Muestra	Tipo de estudio	Fiabilidad			Validez						
					Alpha	r / ICC	Kappa	Estructura	Convergente	Discriminante	Predictiva	Sens.Cambio	Aparente	PC
Wing, JK., Beevor, AS., Curtis, RH., Park, SBG., Hadden, S. and Burns, A. Orrell, M., Yard, P., Handysides, J. and Schapira, R.	1998	UK - Not	2706	Validación/ Fiabilidad	---	(n=100) Tot .86 Sub .82 to .89	---	4 factors	RFS= .65 BPRS= .84	---	---	SI	---	---
		UK - Man				(n=97) Tot .77 Sub .68 to .95								
Bebbington, P., Brugha, T., Hill, T., Marsden, L., and Window S.	1999	UK	100	Validación / Fiabilidad	0.65	0.39	0.57	---	LOCS= .51 GAS= -.40 BPRS= .49 SBS=.59 GHQ= .36 SF-36= -.33	---	---	---	SI	---
Chaplin, R., and Perkins, R.	1999	UK-Lond	57	Validación	---	---	---	---	SBS=.23 SRPS=.24 SCAN=-.12	---	---	---	---	---
		UK-Leic	58						SBS=.44 SRPS=.30 SCAN=-.01					
Wing, JK., Beevor, AS., Curtis, RH., Park, SBG., Hadden, S. and Burns, A.	1999	UK	248	Fiabilidad	---	---	---	---	---	---	---	SI	---	---

Orrell, M., Yard, P., Handsides, J. and Schapira, R.	1999	UK	151	Validación/ Fiabilidad	---	(n=46) Tot= .59 Sub= .47- .85	0.7	---	GAF-d=-.86 (n=129) SCAN=.79 (n=42) SCAN=.44	---	---	---	---	---
Slade, M., Beck, A., Bindman, J., Thornicroft, G., and Wright, S.	1999	UK	259	Dat.Psicometr (VConv CANSAS)	---	---	---	---	CANSAS= .45	---	---	---	---	---
McClelland, R., Trimble, P., Fox, ML., Stevenson MR., and Bell, B.	2000	UK	934	Validación	---	---	---	4 factors	BPRS=.44 GAS=.49	SI	SI	SI	---	---
Audin, K., Margison, F.R., Mellor-Clark, J., Barkham, M.	2001	UK	1688	Sensibilidad al cambio	---	---	---	---	---	---	---	SI	---	---
Slade, M., Cahill, S., Kelsey, W., Powell, R., and Strathdee, G.	2002	UK	605	Dat.Psicometr (VConv TAG)	---	---	---	---	TAG=.71	---	---	---	---	---
Rees, A., Richards, A., and Shapiro, DA.	2004	UK	195	Sensibilidad al cambio	---	---	---	---	---	---	---	SI	---	---
Hunter, R., Mclean, J., Peck, D., Pullen, I., Greenfield, A., Mearthur, W., and Quinn, A., Eaglesham, J., Hagen, S., and Norrie, J.	2004	UK	695	Validación/ Fiabilidad	---	---	0.42	---	---	---	---	SI	SI	---
Salvi, G., Leese, M., and Slade, M.	2005	UK	350	Dat.Psicometr (VConv GAF/TAG)	---	---	---	---	GAF=-.61 TAG=.717 CANSAS= 0.675	---	---	---	---	---
Ecob, R., Croudace, TJ., White, IR., Evans, JE., Harrison, GL., Sharp, D., and Jones, PB.	2006	UK	384	Fiabilidad	---	0.29	---	---	---	---	---	---	---	---

Hansen, L., and Kingdon, D.	2006	UK	40	Dat.Psicometr (VConv TOTI/CPRS)	---	---	---	---	TOTI=.598 CPRS=.794	---	---	---	---	---
Williams, B., Speak, B., Hay P., and Muncer, S.J.	2014	UK	2706	Fiabilidad/ Estructura factorial	0.66	---	---	1 factor	---	---	---	---	---	---
Speak, B., and Muncer, S.	2015	UK	80161	Fiabilidad/ Estructura factorial	0.86	---	---	1 factor	---	---	---	---	---	---
Muncer, S., Bass, M., and Dawkin, M.	2016	UK	989	Estructura factorial	---	---	---	2 factors	---	---	---	---	---	---
Speak, BL., and Muncer, SJ.	2016	UK	80161	Fiabilidad/ Estructura factorial	---	---	---	4 factors	---	---	---	---	---	---
Smith, W., Vitoratou, S., McCrone, P., and Patel, A	2017	UK	12910	Fiabilidad/ Estructura factorial	0.52	---	---	4 factors	---	---	---	---	---	---
Richardson, T., Dasyam, B., Courtney, H., White, L., Tedbury, J., Butt, J. and Newman-Taylor, K.	2019	UK	103	Fiabilidad	0.75	---	---	---	---	---	---	---	---	---
Jones, C.I., Bremner, S., Hooper, R. et al.	2021	UK	950	Sensibilidad al cambio/Protocolo	---	---	---	---	---	---	---	SI	---	---
Eagar, K., Trauer, T., and Mellsop, G.	2005	NZ	12576	Validación/ Fiabilidad	0.8	---	---	4 factors	LSP-16=.68	---	---	---	---	---
Stewart, M.	2009	NZ	25	Validación/ Fiabilidad self- report/others	---	0.72	---	---	GAF=-.65	---	---	---	---	---

[illegible]

[illegible]

Seow, L. L. Y., Page, A. C. and Hooke, G. R.	2020	Australia	102	Fiabilidad	0.65	---	---	---	---	---	---	---	---	---
Parker, S., Arnautovska, U., Siskind, D., Dark, F., McKeon, G., Korman, N., and Harris, M.	2020	Australia	501	RCI	---	---	---	---	---	---	---	---	---	PC=13 (rci=6,14)
Vargas, M.L., Quintanilla, P., Lozano, A., Sendra-Gutiérrez, J.M.	2013	España	74	Dat.Psicometr (VConv ERP)	---	---	---	---	ERP=.36	---	---	---	---	---
Dogmanas, D.	2014	España	402	Validación / Fiabilidad URUGUAY	0.69	---	---	Unidimens ionalidad	---	SI	---	SI	---	PC=7.3 RCI=6.07
Lascorz, D., Lopez, V., Pinedo, C., Trujols, J., Vegué, J., and Perez, V.	2016	España	75	Dat.Psicometr (VConv ENAR)		---	---	---	ENAR-TMG= .474	---	---	---	---	---
Uriarte, J. J., Iglesias, N., Penas, P., Moreno-Calvete, M. C., Álvarez-González, A. and Iraurgi, I.	2024	España	281	Validación / Fiabilidad	0.70	---	---	4 factors	---	---	---	---	---	---
Lovaglio, PG., and Monzani, E.	2011	Italia	2162	Validación / Fiabilidad	T=0.71 .64-.73	---	---	4 factors	---	---	---	SI	---	---
Parabiaghi, A., Barbato, A., D'Avanzo, B., Erlicher, A., and Lora, A.	2005	Italia	9817	Validación / Fiabilidad	.73	---	---	---	---	---	---	---	---	---
Preti, A., Pisano, A., Cascio, MT., Galvan, F., Monzani, E., Meneghelli, A., and Cocchi, A.	2012	Italia	168	Validación / Fiabilidad	.70	FEP N=87 .76	---	---	FEP N=87 BPRS=.78 GAF=-.23	---	NO	SI	---	---
						UHR n=81 .64			UHR n=81 BPRS=.44 GAF=-.35		NO			

Preti, A., Pisano, A., Cascio, MT., Galvan, F., Monzani, E., Meneghelli, A., and Cocchi, A.	2012	Italia	1633	Validación / Estructura factorial -6 ítems	---	---	---	1 factor	---	---	---	---	---
Vita, A., Deste, G., Barlati, S., De Peri, L., Giambra, A., Poli, R., Keefe, RSE., and Sacchetti, E.	2013	Italia	86	Dat.Psicometr (VConv SCORS)	0.78	---	---	---	SCORS=.43	---	---	---	---
Mulder, C.L., Staring, A.B.P., Loos, J., Buwalda, V.J.A., Kuijpers, D., Sytema, S., and Wierdsma A.I.	2004	Países Bajos	559	Validación / Fiabilidad	0.78	0.92 (.82/.97)	---	---	CGI=.51 CANSAS=.71 QOL=-.55 BPRS=.75 GAF=-.59- -.69	---	---	---	---
Tenneij, N., Didden, R., Veltramp, E., and Koot, HM.	2009	Paises Bajos	79	Validación / Fiabilidad	0.48 0.74	0.60	---	---	(n=54) ABCL=.46 (n=52) SRZ-P=-.34 (n=44) ABCL=.59 (n=42) SRZ-P=-.54	---	---	---	---
Buwalda, V.J.A., Draisma, S., Smit, J.H., Swinkels, J.A., Van Tilburg, W.	2011	Paises Bajos	148	Validación / Fiabilidad	0.64	---	---	---	SCL90R=.64	---	---	SI	---
Fassaert, T., Lauriks, S., van de Weerd, S., Theunissen, J., Kikkert, M., Dekker, J., Buster, M., and de Wit, M.	2014	Paises Bajos	107	Dat.Psicometr (VConv SSM-D)	---	---	---	---	SSM-D=-.51	---	---	---	---
Jochems, E.C., Duivenvoorden, H.J., van Dam, A. et al.	2018	Paises Bajos	294	Fiabilidad	0.70	---	---	---	---	---	---	---	---

Albers, W. M. M., Roeg, D. P. K., Nijssen, Y., van Weeghel, J., Bongers, I. M. B.	2018	Países Bajos	395	Fiabilidad	0.78	0.92	---	---	---	---	---	---	---	---
Jansen, L., Hunnik, F., Busschbach, J.J.V., Lijmer, J.G	2019	Países Bajos	50	Fiabilidad	0.63	---	---	---	---	---	---	---	---	PC=14
Kortrijk, H., Schaefer, B., van Weeghel, J., Mulder, C.L., and Kamperman, A.	2019	Países Bajos	2660	Dat.Psicometr (VConv QoL/CANSAS)	---	---	---	---	CNCM= .49 CANSAS= .34	---	---	---	---	---
Kwakernaak, S., Swildens, W.E., van Wel, T.F. et al.	2019	Países Bajos	287	Validación / Fiabilidad	0.75	---	---	---	---	---	---	---	---	---
Liemburg, E., Nienhuis, F. and Veling, W.	2020	Países Bajos	164	Dat.Psicometr (VConv CRDPSS) Póster	---	---	---	---	---	---	---	---	---	---
Sportel, B. E., Aardema, H. and Boonstra, N. et al.	2023	Países Bajos	259	Dat.Psicometr (VConv I.ROC)	---	---	---	---	I.ROC =-.52	---	---	---	---	---
Bartels-Velthuis, A., Ties, K., Visser, E., Arends, J., Pijnenborg, M., Wunderink, L., Jörg, F., Veling, W., Castelein, S., Knegtering, H. and Bruggeman, R.	2020	Países Bajos	156	Dat.Psicometr (VConv BPSQ) Póster	---	---	---	---	BPSQ = .488	---	---	---	---	---
Bech, P., Bille, J., Schüte, T., Søndergaard, S., Waart, S., and Wiese, M.	2003	Denmark	963	Validación / Datos clínicos	---	---	---	---	---	---	---	---	---	---
Øiesvold, T., Bakkejord, T., abd Sexton, JA.	2011	Noruega	118	Validación / Fiabilidad	T=0.67 0.18/ 0.57	---	---	---	SCL90R=.33	---	---	---	---	---

Andreas, S., Harfst, T., Dirmaier, J., Kowski, S., Koch, U., & Schulz, H.	2007	Alemania	543	Fiabilidad	---	.84	---	---	---	---	---	---	---	---
Andreas, S., Harfst, T., Rabung, S., Mestel, R., Schauenburg, H., Hausberg, M., Kowski, S., Koch, U., and Schulz, H.	2009	Alemania	3169	Validación / Fiabilidad	0.62	---	---	---	SCL90R=.55 SF-8=.54 BSS=.60	---	---	SI	---	---
Andreas, S., Harries-Hedder, K., Schwenk, W., Hausberg, M., Koch, U., and Schulz, H.	2010	Alemania	417	Validación / Fiabilidad	0.73	(n=129) .78	---	4 factors	---	---	---	---	---	---
Andreas, S., Rabung, S., Mestel, R., Koch, U., Hausberg, M., and Schulz, H.	2011	Alemania	1918 + 55	Sensibilidad al cambio – Item 8	---	---	---	---	---	---	---	SI	---	---
Theodoridou, A., Jaeger, M., Ketteler, D., Kawohl, W., Lauber, C., Hoff, P., and Rössler, W.	2009	Suiza	100	Dat.Psicmetr (VConv CGI)	---	---	---	---	CGI=.201 AMDP=.352	---	---	SI	---	---
Kocha, N., Wasserfallen, JB., Camusa, D., Gebela, S., Bonsack, C.	2013	Suiza	2414	Sensibilidad al cambio	---	---	---	---	---	---	---	SI	---	---
Egger, ST., Weniger, G., Prinz, S., Vetter, S., and Müller, M.	2015	Suiza	690	Sensibilidad al cambio	---	---	---	---	---	---	---	SI	---	---
Golay, P., Basterrechea, L., Conus, P., and Bonsack, C.	2016	Suiza	2722	Estructura factorial	---	---	---	4 factors	---	---	SI	---	---	---
Böckmann, V., Lay, B., Seifritz, E., Kawohl, W., Roser, P. and Habermeyer, B.	2019	Suiza	554	Validación / Fiabilidad	---	---	---	---	---	---	SI	---	---	---
Egger, S. T., Bobes, J., Theodoridou, A., Seifritz, E. and Vetter, S.	2020	Suiza	30616	Validación / Sensibilidad al cambio	---	---	---	---	---	SI	---	---	---	---

Mötteli, S., Adamus, C., Deb, T., Fröbel, R., Siemerikus, J., Richter, D. and Jäger, M.	2022	Suiza	58	Fiabilidad	0.50	---	---	---	---	---	---	---	---	---
Lauzon, S., Corbière, M., Bonin, JP., Bonsack, C., Lesage, AD., and Ricard, N.	2001	Canadá	194	Validación / Fiabilidad	0.70	0.92	0.81	4 factors	---	---	---	---	---	---
Kisely, S., Campbell, LA., Crossman, D., Gleich, S., and Campbell, J.	2007	Canadá	485	Sensibilidad al cambio	---	---	---	---	---	---	---	SI	---	---
Kisely, S., Campbell, LA., Cartwright, J., Cox, M., and Campbell, J.	2010	Canadá	4620	Validación / Sensibilidad al cambio	---	---	---	---	---	---	SI	SI	---	---
Currie, S.R., Liu, P., Adamyk-Simpson, J. et al.	2019	Canadá	73	Dat.Psicometr (Val Paciente experience)	---	---	---	---	---	---	SI (0.23)	---	---	---
Browne, S., Doran, M., and McGauran, S.	2000	Irlanda	100	Dat.Psicometr (VConv GAF)	---	---	---	---	GAF=-.74	---	---	---	---	---
Mulligan, O., Helmi, L., Adamis, D., O'Mahony, E., McCarthy, G., and MacDonald, A.	2016	Irlanda	280	Dat.Psicometr (VConv GAF)	---	---	---	---	GAF=-.696	---	---	---	---	---
Sweeney, J., Adamis, D., Helmi, L. et al.	2019	Irlanda	470	Dat.Psicometr (VConv GAF)	---	---	---	---	GAF=-.919	---	---	---	---	---
Pěč, O., Čechová, D., Pěčová, J., Dragomirecká, E., Šelepová, P., Probstová, V., and Bubela, M.	2009	Rep. Checa	40	Validación / Fiabilidad	0.79	T=.908 .602-.794	---	---	GAF=-.72	---	---	---	---	---
Machová, K., Juríčkova, V., Kasparová, A., Petrova, K., Slađkova, B. and Svobodova, I.	2023	Rep. Checa	57	Validación / Fiabilidad Sensibilidad al cambio	0.52	---	---	---	---	---	---	SI	---	---

Drakonakis, N., Stylianidis, S., Peppou, L.E. et al.	2021	Grecia	544	Sensibilidad al cambio	---	---	---	---	---	---	---	SI	---	---
Anjara, S.G., Bonetto, C., Ganguli, P., Setiyawati, D., Mahendradhata, Y., Yoga, B. H., et al.	2019	Indonesia	287	Validacion/fiabilidad	---	---	---	---	---	---	---	SI	---	---

Anexo 5: Hoja informativa del proyecto

HOJA DE INFORMACIÓN AL USUARIO PROYECTO REE BIZKAIA

Estimado Sr/Sra.

En los últimos años la participación de los usuarios como agentes activos en su proceso de atención se ha convertido en un elemento fundamental para afrontar y mejorar la atención de las personas que padecen enfermedades crónicas.

Para la Red de Salud Mental de Bizkaia es un objetivo importante conocer qué aspectos son importantes para las personas afectadas en su proceso de recuperación y qué barreras se encuentran, incluyendo el estigma. Por ello desde la Red de Salud Mental de Bizkaia queremos llevar a cabo un estudio que nos permita conseguir información sobre estos aspectos.

Usted, como persona afectada y que recibe atención en nuestros servicios, reúne los requisitos necesarios para participar en el estudio, por lo que solicitamos su colaboración para poder recoger la información necesaria para llevarlo a cabo. La información necesaria se obtendrá de su historia clínica de forma absolutamente confidencial, de manera que sus datos no podrán ser relacionados de ninguna manera con usted. Además, se le solicitará una entrevista personal en la que un entrevistador aplique el cuestionario al que nos gustaría que respondiese.

NO se realizaría ninguna exploración de tipo invasivo, ni este estudio modifica de ninguna manera el tratamiento que está usted recibiendo en la actualidad.

El estudio no supone ningún riesgo, y en cambio el conocimiento que se puede obtener en este estudio de aquellos aspectos que son importantes para usted y en qué medida nuestros servicios pueden ayudarle mejor en su proceso de recuperación puede ser muy importante para poder dar una atención de mayor calidad y adaptada a sus necesidades.

Se garantiza la confidencialidad de los datos personales de los usuarios de acuerdo con la Ley Orgánica 15/1999 de 13 de diciembre, de protección de datos de carácter personal.

De acuerdo con la legislación mencionada, usted puede ejercer los derechos de acceso, modificación, oposición y cancelación de datos, para lo cual se deberá dirigir al investigador de estudio.

La participación en el estudio es voluntaria, pudiendo retirarse del mismo en cualquier momento. En cualquier caso, tanto si se decide participar en el estudio como si no, la atención que reciba posteriormente no se verá afectada en absoluto.

Los datos recogidos para el estudio estarán identificados mediante un código y sólo el investigador podrá relacionar dichos datos con el usuario, por lo tanto, la identidad no será revelada.

Si tiene cualquier duda sobre el estudio el equipo que le atiende, y el propio equipo investigador está a su disposición para cualquier aclaración

Puede dirigir sus preguntas, dudas o sugerencias a: booms.investigacion@osakidetza.net

Anexo 6: Consentimiento informado para el usuario participante

CONSENTIMIENTO INFORMADO DEL USUARIO

Yo (Nombre y apellidos)con DNI... ,

He leído la información que se me ha entregado. He podido hacer preguntas sobre el estudio.

He recibido suficiente información sobre el estudio. He hablado con (profesional de referencia) Comprendo que la participación es voluntaria.

Comprendo que me puedo retirar del estudio:

1. Cuando quiera
2. Sin tener que dar explicaciones
3. Sin que esto repercuta en mis cuidados médicos

También he sido informado de forma clara, precisa y suficiente de los siguientes extremos que afectan a los datos personales que se contienen este consentimiento.

Estos datos serán tratados y custodiados con respeto a mi intimidad y a la vigente normativa de protección de datos.

Sobre estos datos me asisten los derechos de acceso, rectificación, cancelación y oposición en la dirección de contacto que figura en este documento.

Estos datos no podrán ser cedidos sin mi consentimiento expreso.

Declaro que he leído y conozco el contenido del presente documento. Y por ello, firmo este consentimiento informado de forma voluntaria para MANIFESTAR MI DESEO DE PARTICIPAR EN ESTE ESTUDIO SOBRE LA EVALUACIÓN DE LA PERCEPCIÓN SUBJETIVA DE RECUPERACIÓN. EVALUACIÓN ORIENTACIÓN DE SERVICIOS DE SALUD MENTAL DE LA RSMB A LA RECUPERACIÓN. Proyecto DREEM-Bizkaia, hasta que decida lo contrario. Al firmar este documento no renuncio a ninguno de mis derechos. Recibiré una copia de este consentimiento para guardarlo y poder consultarlo en el futuro.

Fecha:

Fdo.:

Anexo 7: Hoja informativa del Proyecto para tutor legal

HOJA DE INFORMACIÓN AL TUTOR LEGAL DEL USUARIO PROYECTO REE BIZKAIA

Estimado Sr/Sra.

En los últimos años la participación de los usuarios como agentes activos en su proceso de atención se ha convertido en un elemento fundamental para afrontar y mejorar la atención de las personas que padecen enfermedades crónicas.

Para la Red de Salud Mental de Bizkaia es un objetivo importante conocer qué aspectos son importantes para las personas afectadas en su proceso de recuperación y qué barreras se encuentran, incluyendo el estigma. Por ello desde la Red de Salud Mental de Bizkaia queremos llevar a cabo un estudio que nos permita conseguir información sobre estos aspectos.

Su familiar/persona tutelada, como persona afectada y que recibe atención en nuestros servicios, reúne los requisitos necesarios para participar en el estudio, por lo que solicitamos su colaboración para poder recoger la información necesaria para llevarlo a cabo. La información necesaria se obtendrá de su historia clínica de forma absolutamente confidencial, de manera que sus datos no podrán ser relacionados de ninguna manera con la persona. Además, se le solicitará una entrevista personal en la que un entrevistador le pasará un cuestionario para evaluar su percepción sobre la recuperación (cuestionario DREEM), así como un cuestionario sobre el estigma (cuestionario ISMI) y un cuestionario sobre percepción de la salud (cuestionario Euro-Qol5d).

NO se realizaría ninguna exploración de tipo invasivo, ni este estudio modifica de ninguna manera el tratamiento que está usted recibiendo en la actualidad.

El estudio no supone ningún riesgo, y en cambio el conocimiento que se puede obtener en este estudio de aquellos aspectos que son importantes para las personas afectadas y en qué medida nuestros servicios pueden ayudarles mejor en su proceso de recuperación puede ser muy importante para poder dar una atención de mayor calidad y adaptada a sus necesidades.

Se garantiza la confidencialidad de los datos personales de los usuarios de acuerdo con la Ley Orgánica 15/1999 de 13 de diciembre, de protección de datos de carácter personal.

De acuerdo con la legislación mencionada, usted puede ejercer los derechos de acceso, modificación, oposición y cancelación de datos, para lo cual se deberá dirigir al investigador de estudio.

La participación en el estudio es voluntaria, pudiendo retirarse del mismo en cualquier momento. En cualquier caso, tanto si se decide participar en el estudio como si no, la atención que reciba posteriormente no se verá afectada en absoluto.

Los datos recogidos para el estudio estarán identificados mediante un código y sólo el investigador podrá relacionar dichos datos con el usuario, por lo tanto, la identidad no será revelada.

Si tiene cualquier duda sobre el estudio el equipo que le atiende, y el propio equipo investigador está a su disposición para cualquier aclaración

Puede dirigir sus preguntas, dudas o sugerencias a: booms.investigacion@osakidetza.net

Anexo 8: Consentimiento informado en caso de tutor legal

CONSENTIMIENTO INFORMADO AL TUTOR LEGAL DEL USUARIO

Yo (Nombre y apellidos)con DNI ,

He leído la información que se me ha entregado. He podido hacer preguntas sobre el estudio.

He recibido suficiente información sobre el estudio. He hablado con (Profesional de referencia) Comprendo que la participación es voluntaria.

Comprendo que me puedo retirar del estudio:

1. Cuando quiera
2. Sin tener que dar explicaciones
3. Sin que esto repercuta en mis cuidados médicos

También he sido informado de forma clara, precisa y suficiente de los siguientes extremos que afectan a los datos personales que se contienen este consentimiento.

Estos datos serán tratados y custodiados con respeto a mi intimidad y a la vigente normativa de protección de datos. Sobre estos datos me asisten los derechos de acceso, rectificación, cancelación y oposición en la dirección de contacto que figura en este documento. Estos datos no podrán ser cedidos sin mi consentimiento expreso.

Declaro que he leído y conozco el contenido del presente documento. Y por ello, firmo este consentimiento informado de forma voluntaria para MANIFESTAR MI DESEO DE QUE.....PARTICIPE EN ESTE ESTUDIO LA EVALUACIÓN DE LA PERCEPCIÓN SUBJETIVA DE RECUPERACIÓN. EVALUACIÓN ORIENTACIÓN DE SERVICIOS DE SALUD MENTAL DE LA RSMB A LA RECUPERACIÓN. Proyecto

REE-Bizkaia, hasta que decida lo contrario. Al firmar este documento no renuncio a ninguno de mis derechos ni los de mi familiar. Recibiré una copia de este consentimiento para guardarlo y poder consultarlo en el futuro.

Fecha:

Fdo.:

Anexo 9: Evidencias de Adecuación Psicométrica de la Versión Española de la Escala HoNOS

Recibido / Recebido: 25.05.2022 - Aceptado / Aceite: 14.01.2024

<https://doi.org/10.21865/RIDEP72.2.10>

Evidencias de Adecuación Psicométrica de la Versión Española de la Escala HoNOS

Evidence of the Psychometric Adequacy of the Spanish Version of the HoNOS

José Juan Uriarte¹, Nerea Iglesias², Patricia Penas³, María-Concepción Moreno-Calvete⁴, Alexander Álvarez-González⁵ y Ioseba Iraurgi⁶

Resumen

La evaluación de resultados en programas de intervención en salud mental se presenta como una alternativa de elección para la valoración de la efectividad de los tratamientos. Para su implementación se hace necesario contar con instrumentos de medida adecuadamente adaptados y validados al contexto de aplicación. El objetivo del estudio es ofrecer evidencias de la adecuación psicométrica del HoNOS en su adaptación al castellano. Participan en el estudio 281 usuarios seleccionados por muestreo probabilístico-estratificado. Se utiliza la escala HoNOS y se contrasta mediante técnicas de análisis factorial confirmatorio la estructura dimensional que mejor ajusta los datos al modelo teórico de agrupación de síntomas. Se presentan evidencias que muestran la escala HoNOS como un instrumento adecuado en la valoración de resultados en enfermedad mental grave. Los resultados permiten proponer tanto la utilización de un índice global de valoración, como el uso de las dimensiones específicas propuestas por los autores del instrumento.

Palabras clave: HoNOS, validación, propiedades psicométricas, análisis factorial

Summary

Outcome evaluation in mental health intervention programs is presented as an alternative of choice for assessing the effectiveness of treatments. For its implementation, it is necessary to have measuring instruments adequately adapted and validated to the context of application. The aim of the study is to provide evidence of the psychometric adequacy of HoNOS adapted to Spanish. A total of 281 users selected by probabilistic-stratified sampling participated in the study. The HoNOS scale was used and the dimensional structure that best fits the data to the theoretical model of symptom clustering was tested using confirmatory factor analysis techniques. Evidence was presented showing the HoNOS scale as an appropriate instrument for result assessment in severe mental illnesses. These results make it possible to propose both the use of an overall assessment index and the use of the specific dimensions proposed by the instrument's authors.

Keywords: HoNOS, validity, psychometric properties, factorial analysis

¹Jefe UGC. Instituto de Investigación Sanitaria Biobizkaia, Osakidetza, Red de Salud Mental de Bizkaia. Correo: josejuan.uriarteuriarte@osakidetza.eus

²Universidad de Deusto. Facultad de Ciencias de la Salud. Dpto de Psicología. Correo: nereai Iglesias@deusto.es

³Universidad de Deusto. Facultad de Ciencias de la Salud. Dpto de Psicología. Correo: patricia.penas@deusto.es

⁴Instituto de Investigación Sanitaria Biobizkaia, Osakidetza, Red de Salud Mental de Bizkaia. Correo:

mariaconcepcion.morenocalvete@osakidetza.eus

⁵Universidad de Deusto. Facultad de Ciencias de la Salud. Dpto de Psicología. Correo: alvarez@opendeusto.es

⁶Universidad de Deusto. Facultad de Ciencias de la Salud. Dpto de Psicología. DeustoPsych. Investigador Principal del Grupo "Evaluación, Clínica y Salud". Correo: ioseba.iraurgi@deusto.es (Autor de correspondencia)

Revista Iberoamericana de Diagnóstico y Evaluación - e Avaliação Psicológica, RIDEP - Nº72 - Vol. 2 - 145-156- 2024
ISSN: 1135-3848 print / 2183-6051 online

This work is licensed under CC BY-NC 4.0. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc/4.0/>

Introducción

La evaluación de resultados en salud mental, al igual que en otras disciplinas sanitarias, se ha desarrollado de forma creciente como opción válida en la valoración de la efectividad de las intervenciones clínicas y como alternativa a la evaluación de la eficacia que ofrecen los ensayos clínicos (Slade, 2002; Schmitter-Edgecombe et al., 2012). En el contexto natural de intervención clínica la utilización de procedimientos de investigación controlados es una opción de difícil implementación, de ahí que la metodología desarrollada por la evaluación de resultados – basada en un procedimiento de aplicación sistemático que permita la estimación del cambio (comparaciones intra y entre grupos), la utilización de instrumentos de medida adecuadamente ajustados y un análisis correcto de los datos– sea una opción cada vez más utilizada (Wise, 2004; Walton et al., 2015).

La utilización de instrumentos psicométricos en la evaluación de resultados en salud mental es una opción que ha venido utilizándose durante décadas. Existe todo un repertorio de instrumentos y escalas de medida (Slade, 2002; McDowell, 2006) orientados a la valoración de signos y síntomas psíquicos específicos (depresión –BDI, HDS– o alexitimia –TAS-20–, por ejemplo) o a la valoración de una generalidad de manifestaciones valorados a través de instrumentos multicomponentes (SCL-90, GHQ, OQ-45, CORE-ON y/o SF36). Para su correcta utilización en el ámbito de evaluación de resultados resulta imprescindible que dichos instrumentos se hayan adaptado adecuadamente al idioma y cultura en la que se utilicen (Wild et al., 2009; Mokkink et al., 2010) y sea probada su capacidad y adecuación de medida en los términos clásicos de validez y fiabilidad (Scholtes et al., 2001).

Uno de los instrumentos más utilizados a nivel internacional, y más versátiles en su aplicación clínica en el trastorno mental grave, es el HoNOS (Wing et al., 1998). En el año 1992, la cartera de salud el Gobierno Británico interesada en buscar métodos eficientes para la toma de decisiones edita un libro blanco (Department of Health, 1992) con el objetivo de mejorar la salud y el bienestar social de las personas con enfermedad mental, que se podría entender como una mejora clínica y social,

así como preservar la funcionalidad con el fin de prevenir el deterioro. Para evaluar este propósito, el Real Colegio de Psiquiatras de Reino Unido desarrolló el instrumento Health of the Nation Outcome Scales - HoNOS (Wing et al., 1998). Se trata de una herramienta clínicamente aceptada y de uso fácil para ser completada por el profesional sanitario; cubre un amplio espectro de problemas clínicos y disfunciones sociales; posee una fiabilidad aceptable y es concurrente con otros instrumentos ya validados; y además es sensible al cambio o a la ausencia del mismo a lo largo del tiempo (Wing et al., 2000; Pirkis et al., 2005; Iglesias et al., 2019).

En su propuesta final original (Wing et al., 1998), los autores proponen un instrumento de 12 ítems que explorarían cuatro dimensiones en las que se manifiestan los trastornos mentales graves: problemas conductuales, deterioro, problemas clínicos y problemas sociales. La investigación posterior desarrollada por otros autores sobre el HoNOS ha mostrado otras alternativas de clasificación de los ítems (Tabla 1) que, respondiendo a su configuración matemática obtenida mediante técnicas de análisis factorial, ofrece una amalgama de posibles configuraciones conceptuales de las manifestaciones de resultados en trastorno mental grave: Trauer (1999) propone el uso de cinco factores, otros autores (Preston, 2000; McClelland et al., 2000; Newnham et al., 2009; Speak & Muncer, 2016) encuentran cuatro factores correlacionados distintos al original, Lovaglio y Monzani (2012) un único factor con seis ítems y Speak y Muncer (2015) un factor con cuatro ítems que hacen referencia al factor social del modelo original. En la base de todas estas discrepancias se halla el propio comportamiento de los ítems que ofrecen una correlación moderada-baja entre los mismos (Trauer, 1999).

En su elección inicial (Wing et al., 1998) se optó por una escala breve que reflejase todo el abanico de manifestaciones en el ámbito del trastorno mental grave, imponiéndose las consideraciones prácticas de medir el fenómeno sobre las matemáticas. No obstante, en los últimos años se ha venido promoviendo una alternativa en la valoración de la configuración de medida de los instrumentos basada en el análisis bifactorial (Reise et al., 2013; Rodríguez & Reise, 2016) que pudiera ser una alternativa en la comprensión

Tabla 1. Estructuras factoriales sugeridas para el HoNOS

Wing (1998)	Ítems	Trauer (1999)	Preston (2000)	McClelland (2000)	Newnham (2009)	Lovaglio (2012)	Speak (2015)	Speak (2016)
F1 –	It01 – Agresividad	1	1	1	1			4
Problemas	It02 – Autoagresiones	3	1	2	1			2
conductuales	It03 – Uso de sustancias	1	1	1	1			3
F2 –	It04 – Disfunción cognitiva	2	2	1	2	1		1
Deterioro	It05 – Discapacidad Física	2	2	3	2			1
F3 –	It06 – Depresión	3	3	1	3	1		4
Problemas	It07 – Alucinaciones/Delirios	5	3	2	2			2
Clinicos	It08 – Otros Síntomas	3	3	4	3			2
	It09 – Relaciones Sociales	3-4	4	1	1	1	1	3
F4 –	It10 – Funcionamiento general	4	4	1	2	1	1	1
Problemas	It11 – Problemas de residencia	4	4	1	4	1	1	3
sociales	It12 – Problemas ocupacionales	4	4	1	4	1	1	1-3

Nota. Los factores de Preston, McClelland mantienen las denominaciones propuestas por Wing en el estudio original. Trauer define los factores de la siguiente forma: F1=Conducta, F2=deterioro, F3=Depresión, F4=Problemas sociales y F5=Alucinaciones. Newnham los define: F1=Problemas conductuales o antisociales, F2=Deterioro, F3=Síntomas y F4=Socio-económico. Speak (2016) los define: F1=Bienestar personal, F2=Bienestar emocional, F3=Bienestar social y F4=Alteraciones severas.

del comportamiento psicométrico del HoNOS.

El uso del HoNOS en España se ha posibilitado a través de la adaptación realizada por el grupo de Uriarte en colaboración con los autores del instrumento (Uriarte et al., 1999). Si bien ha sido utilizado en distintos estudios (Gisbert & Cid-Colom, 2010; Frades et al., 2011; Osta-Samanes et al., 2012; Garay-Arostegui et al., 2014; Escalada et al., 2014, 2015a, 2015b; Ribas et al., 2016) y a excepción del estudio de validación en población uruguaya (Dogmanas, 2014) y de una adaptación realizada para personas con discapacidad de aprendizaje (Esteba-Castillo, et al., 2018), a día de hoy no se conoce ningún estudio que informe sobre sus características de medida y adecuación psicométrica al español. A este respecto, el objetivo del presente estudio es ofrecer datos del comportamiento psicométrico del HoNOS en una muestra representativa de la población atendida en la Red de Salud Mental de Bizkaia (RSMB) y delimitar la estructura dimensional del instrumento para la obtención de índices de medida que permitan la evaluación de resultados en enfermedad mental grave.

Método

Muestra

En el contexto de un estudio de evaluación de la recuperación de personas con enfermedad mental grave en tratamiento en la Red de Salud Mental de Bizkaia (RSMB), se obtuvo una muestra representativa de 281 participantes seleccionados de un total de 1949 usuarios censados en la RSMB asumiéndose un error de estimación de 5.1% para

un nivel de confianza del 95%. Para la selección se llevó a cabo una estratificación en función del sexo, edad y tipo de centro asistencial con asignación aleatoria a cada estrato. Como criterios de inclusión se consideraron: ser mayor de 18 años y la pertenencia a la RSMB. Como criterios de exclusión: la no obtención del consentimiento informado, dificultades en la comunicación y presentar un estado clínico que no permita la colaboración del o la participante.

Los participantes presentaban una edad media de 48.8 años (DE=11.2) con edades comprendidas entre los 23 y 81 años, siendo preferentemente hombres (n=174; 61.9%). Respecto a su distribución entre los distintos programas de tratamiento, 164 recibían asistencia en los Centros de Salud Mental (CSM), 74 en Hospitales de Día (HD), 22 en el Tratamiento Asertivo Comunitario (TAC) y 21 en Hospitales Psiquiátricos (HP). El diagnóstico más frecuente fue de esquizofrenia (54.2%), seguido del trastorno bipolar (12.3%) y el trastorno esquizoafectivo (9.9%), con una media de 17.85 años en tratamiento (DE=8.64, rango=2-45). Un 76.9% de los participantes han sufrido alguna hospitalización a lo largo de su evolución con una media de 5.16 hospitalizaciones (DE=7.89). Un 78.5% de los casos tenían acreditado un grado de minusvalía (media=65.06, DE=9.82), asimismo un 77% no tenían reconocido ningún grado de dependencia. En relación a la convivencia, un 66.1% viven con sus familias, un 1.4% autónomamente y un 7.5% en alojamientos supervisados.

Instrumentos

Health of the Nation Outcome Scale – HoNOS (Wing et al., 1998). Se trata de una escala de 12

ítems de aplicación por parte del personal clínico que a partir de una escala de intensidad de 0 a 4 puntos (0=sin problema, 1=problema menor que no requiere intervención, 2=problema leve pero claramente presente, 3=problema de moderada gravedad y 4=problema grave o muy grave) se valora la gravedad del problema que tiene el paciente respecto a cada uno de los ítems. Permite obtener una puntuación global a partir de la contribución de todos los ítems, y cuatro medidas específicas correspondientes a las dimensiones teóricas propuestas por los autores. Existe también la posibilidad de utilizar cada ítem como un indicador único de la manifestación a la que hace referencia (Bebbington et al., 1999), si bien esta opción es poco utilizada y poco recomendada (Speak & Muncer, 2015). Un reciente estudio de revisión sistemática de las evidencias de validez del HoNOS (Iglesias et al., 2019) ofrece datos de su consistencia y valores de precisión de medida (fiabilidad) con coeficientes que oscilan entre .37 y .86, siendo la fiabilidad promedio integrada por meta-análisis de .81. En el presente estudio se ha utilizado la adaptación al castellano utilizada por el grupo de Uriarte y colaboradores (1999).

Escala de Evaluación de la Actividad Global – EEAG (Endicott et al., 1976; Bobes et al., 2002). Es una escala de valoración de 0 a 100 puntos diseñada para que un clínico evalúe el nivel general de actividad de un paciente, en la cual una mayor puntuación significaría un mejor funcionamiento. Este instrumento constituye la información del eje V del diagnóstico multiaxial del DSM (APA, 2004).

Escala de Impresión Clínica Global – ICG (Guy, 1976; 2004), en concreto la sub-escala que evalúa la gravedad del cuadro clínico (CGI-SI), se trata de una escala Likert donde el clínico asigna un valor de 0 (no evaluado) a 7 (extremadamente enfermo) donde se califica la gravedad global del padecimiento del paciente.

Escala analógico visual del cuestionario de Salud Euro-QoL5D – EQ5D (EuroQol Group, 1990; Herdman et al., 2001). Se trata de una escala analógico-visual en forma de termómetro donde el participante marca “su estado de salud hoy”, siendo 0 el peor estado de salud que pueda imaginarse y 100 el mejor.

Procedimiento

De manera protocolizada, el programa de Trastorno Mental Grave de la RSMB tiene implementado un procedimiento de valoración del estado de gravedad y evolución en el programa de los usuarios que acceden al mismo. La valoración de cada sujeto en el momento de evaluación que le corresponde es realizada de forma colegiada por el equipo que le asiste a través de los instrumentos HoNOS, ICG y EEAG, habiéndose solicitado al participante la cumplimentación del Euro-QoL5D. Para el presente estudio se tomó como base de análisis la última valoración realizada de cada participante seleccionado para el estudio. En todo momento el procedimiento de evaluación sigue los estándares de comportamiento ético en la investigación y cuenta con la aprobación del CEIC (Comité Ético de Investigación Clínica de Euskadi - Osakidetza).

Análisis de datos

Para el análisis de los ítems de HoNOS se utilizaron la frecuencia (n), porcentaje, media (M), desviación estándar (DE), asimetría (As), correlación entre el ítem con el resto de la escala (r), Alfa de Cronbach (α) si se elimina elemento y Alfa de Cronbach para el total y las dimensiones. Para el análisis de asociación entre variables se utilizó el coeficiente de correlación de Pearson (r).

Un primer análisis factorial exploratorio se llevó a cabo con el programa Factor 10 (Lorenzo-Seva & Ferrando, 2019). Se verificó la idoneidad de la matriz de correlación a través de la prueba Kaiser-Meyer-Olkin y la prueba de esfericidad de Bartlett. El test de Análisis Paralelos (PA; Parallel Analysis) y el método de Hull se utilizaron como criterios de extracción para el número de factores. Asimismo, se analizó la normalidad multivariada con la prueba de Mardia. Los análisis factoriales confirmatorios (CFA) se llevaron a cabo con el programa EQS 6.1 (Bentler, 2006), utilizando el método de estimación de máxima verosimilitud. Se probaron todos aquellos modelos encontrados en la revisión teórica, incluyendo las estructuras bifactoriales de estos. Para evaluar el nivel de bondad de los distintos modelos se han utilizado los siguientes indicadores: la prueba chi-cuadrado χ^2 de igualdad entre matriz y grados de libertad (g.l.), el criterio de información de Akaike (AIC), bondad del índice de ajuste (GFI), el índice

Tabla 2. Estadísticos descriptivos y medidas de tendencia central de los ítems y dimensiones del HoNOS

Ítems	Puntuaciones					Estadísticos Descriptivos			Fiabilidad		
	0	1	2	3	4	M	DE	CV	As	r	h ²
1	74.0	13.5	9.3	2.5	.7	.42	.82	195.24	2.14	.32	.69
2	96.1	1.1	1.1	.7	1.1	.10	.53	530.00	6.07	.32	.69
3	76.5	10.0	8.9	3.9	.7	.42	.86	204.76	2.05	.20	.70
4	29.9	35.2	26.7	8.2	.0	1.13	.94	83.19	.34	.27	.69
5	39.1	21.0	25.6	11.4	2.8	1.18	1.15	97.46	.55	.17	.71
6	41.3	22.8	24.6	9.6	1.8	1.08	1.10	101.85	.62	.22	.70
7	43.1	27.8	21.0	7.1	1.1	.95	1.01	106.32	.76	.31	.69
8	43.1	14.6	24.2	15.7	2.5	1.20	1.22	101.67	.47	.38	.68
9	13.9	20.6	38.4	23.8	3.2	1.82	1.05	57.69	-.19	.56	.65
10	28.1	23.1	30.6	16.7	1.4	1.40	1.11	79.29	.15	.65	.63
11	66.5	12.5	11.4	3.9	5.7	.70	1.17	167.14	1.63	.34	.69
12	54.8	16.4	17.4	6.8	4.6	.90	1.19	132.22	1.12	.38	.68
Dimensiones											
Problemas Conductuales						.94	1.49	158.51	1.81	.38	.43
Deterioro						2.31	1.52	65.80	.20	.30	.29
Problemas Clínicos						3.23	2.31	71.52	.72	.37	.45
Problemas Sociales						4.82	3.21	66.19	.69	.54	.68
Total						11.30	5.98	52.92	.55		.70

Nota. 0=Sin problema, 1=problema menor que no requiere intervención, 2=problema leve pero claramente presente, 3=problema de moderada gravedad o 4=problema grave o muy grave; M=media, DE=Desviación estándar, CV=Coefficiente de Variación, As=Asimetría, α =Alfa de Cronbach si se elimina el elemento y para los totales de la escala se informa de alfa de Cronbach obtenida, h²=Comunalidad.

ajustado no normado (NNFI), el índice de ajuste comparativo (CFI), el residuo cuadrático medio estandarizado (SMSR), el error cuadrático medio de aproximación (RMSRA) junto con su intervalo de confianza (LI-LS) del 90%. Se utilizaron los siguientes criterios para evaluar la bondad de ajuste: los valores de GFI, NNFI y CFI deben superar el valor de .90 (cuanto mayor sea el valor, mayor será el ajuste) y los valores de SMSR y RMSEA deberán ser menores o iguales a .05, aunque también pueden asumirse valores de hasta .07.

Para el contraste de medias se utilizaron pruebas de análisis de varianza (F), que en caso de falta de homocedasticidad entre grupos se aplicó la corrección del método robusto basado en la prueba de Brown-Forsythe. En caso de significación estadística de la prueba F, el contraste de medias entre grupos se realizó mediante pruebas post-hoc basados en el método de Scheffé.

Resultados

En la Tabla 2 se presentan los resultados descriptivos de los ítems de la escala HoNOS y sus características de consistencia interna. Cinco de los doce ítems presentan efectos suelo notorios con porcentajes de respuesta superiores al 50% en la respuesta de ausencia del sintoma (valor 0) y, portanto, con asimetrías positivas superiores al valor unidad. El ítem 9, por el contrario, es el único

que presenta asimetría negativa, pero con un valor dentro de la normalidad ($As=-.19$).

La correlación promedio del ítem con el total de la escala es moderada ($r=.34$), con valores que oscilan entre .17 y .65. Por su parte, la consistencia interna para el total de la escala es alta ($\alpha=.70$), y la retirada de ninguno de sus componentes logra mejorar sustantivamente este valor de fiabilidad. No obstante, si se atiende a la configuración dimensional propuesta por Wing y colaboradores (1998) los valores de consistencia interna resultan bajos (valores α entre .10 y .45), exceptuando la dimensión de problemas sociales que ofrece un coeficiente ($\alpha=.68$) próximo a la consideración de alto.

El análisis de la matriz de correlaciones realizada con el programa FACTOR presenta un valor del coeficiente de Mardia de 65.12, indicando la presencia de asimetría multivariada, por lo que se sugiere la utilización de matrices policóricas en el análisis de la estructura factorial de los ítems. Asimismo, la prueba de Análisis Paralelos (Pararell Analysis) y el Método Hull utilizados para la extracción de factores sugieren la retención de un único factor. En la Tabla 1 se han presentado las distintas estructuras factoriales sugeridas en la literatura para el HoNOS. De la propuesta original de cuatro factores de Wing et al (1998) a otras alternativas también de cuatro factores (Preston, 2000, McClelland et al., 2000; Newnhan et al., 2009) o de cinco factores (Trauer, 1999), e.

Tabla 3. Índices de los Análisis Factoriales Confirmatorios de los distintos modelos estructurales (índices robustos)

Modelo	χ^2	gl	χ^2 / gl	AIC	GFI	NNFI	CFI	SMR	RMSEA	LI	LS
Un factor	126.95	54	2.35	18.96	.96	.85	.88	.069	.069	.054	.085
Wing (4-factores)	382.96	54	7.09	274.96	.83	.32	.45	.150	.148	.133	.161
Wing (4-factores correlacionados)	100.96	48	2.10	4.96	.97	.88	.91	.064	.063	.045	.080
Trauer (5-factores cor.)	79.68	43	1.85	-6.32	.98	.91	.94	.052	.055	.036	.074
McClelland (4-factores cor)	93.12	48	1.94	28.13	.93	.80	.85	.063	.058	.040	.075
Newnham (4-factores cor.)	82.41	45	1.83	-7.59	.98	.91	.94	.061	.054	.035	.073
Lovaglio	40.72	9	4.52	22.73	.98	.79	.88	.064	.112	.078	.148
Speak (4 factores cor)	96.68	47	2.05	2.68	.97	.88	.92	.063	.061	.044	.079
Bifactorial – Wing	50.43	36	1.40	-21.56	.99	.95	.97	.045	.038	.000	.061
Bifactorial – McClelland	60.14	36	1.67	-11.86	.95	.85	.92	.048	.049	.026	.070
Bifactorial – Newnham	44.54	36	1.24	-27.45	.99	.97	.98	.041	.029	.000	.054
Bifactorial – Speak	52.31	35	1.49	-17.69	.99	.95	.97	.046	.042	.013	.064

Nota. χ^2 : Prueba Ji cuadrado de Bondad de Ajuste; gl: Grados de Libertad; AIC: Criterio de Información de Akaike; GFI: Índice de Bondad Global; NNFI: Índice de Ajuste No Normado; CFI: Índice de Ajuste Comparativo; SMR: Residuo Cuadrático Medio Estandarizado; RMSEA: Error Cuadrático Medio de Aproximación; LI - LS: Limite de Confianza Inferior - Superior del 90% del RMSEA.

Tabla 4. Validez concurrente

	HoNOS			
	Problemas conductuales	Deterioro	Problemas clínicos	Problemas sociales
EEAG – Evaluación Actividad Global	-.276**	-.305**	-.453**	-.484**
ICG – Impresión Clínica Global	.270**	.332**	.436**	.482**
EuroQol – Calidad de Vida	-.044	-.019	-.300**	-.123*
HoNOS				
Problemas conductuales	1			
Deterioro	.126*	1		
Problemas clínicos	.271**	.100	1	
Problemas sociales	.362**	.362**	.368**	1

Nota. * $p < .05$ y ** $p < .001$.

incluso de un único factor con restricción de ítems a un número de seis (Lovaglio & Monzani, 2012) o de cuatro elementos (Speak & Muncer, 2015). Las diferentes propuestas dimensionales han sido probadas mediante técnicas de análisis factorial confirmatorio en la configuración propuesta por cada autor, así como en el caso de dichas configuraciones bajo el supuesto de un modelo bifactorial de un factor general y las correspondientes dimensiones propuestas en cada modelo. En la Tabla 3 se presentan los índices de ajuste de cada uno de los modelos sometidos a prueba. El modelo que ha mostrado unos peores índices de ajuste ha sido el propuesto originalmente por Wing, mejorando los resultados cuando los cuatro factores se han correlacionado. Teniendo en cuenta los índices robustos y los cálculos de la invarianza entre modelos, podemos observar cómo las estructuras bifactoriales propuestas por los distintos autores ajustan mejor. El modelo bifactorial con cuatro factores correlacionados de Wing et al (1998) [$\chi^2(36)=50.42$, $p < .001$, AIC=-21.56, CFI=.97, NNFI=.95, RMSEA=.038 (.000-.061)], el bifactorial de cuatro factores correlacionados de McClelland et al. (2000) [$\chi^2(36)=60.14$, $p < .001$, AIC=-11.86, CFI=.92,

NNFI=.82, RMSEA=.049 (.026-.070)], el bifactorial de cuatro factores correlacionados de Newnham et al (2009) [$\chi^2(36)=44.54$, $p < .001$, AIC=-27.45, CFI=.98, NNFI=.97, RMSEA=.029 (.000-.054)] y el modelo bifactorial de cuatro factores correlacionados de Speak et al. (2016) [$\chi^2(35)=52.31$, $p < .001$, AIC=-17.69, CFI=.97, NNFI=.95, RMSEA=.042 (.013-.064)], obteniendo unos índices adecuados.

Los datos de la validez concurrente de la escala HoNOS se presentan en la Tabla 4. Las distintas dimensiones del HoNOS correlacionan significativamente y de forma congruente con los índices de valoración clínica: de forma negativa con el EEAG (una mayor gravedad en las escalas HoNOS se asocian a una peor funcionalidad reflejada por bajas puntuaciones en la EEAG) y de forma positiva con el ICG (una mayor puntuación en ambos instrumentos indica una mayor gravedad o severidad sintomática). Por su parte, los valores de correlación de la calidad de vida percibida por el participante respecto a las valoraciones HoNOS del clínico, sólo han resultado estadísticamente significativas para las dimensiones de problemas clínicos ($r=-.30$) y problemas sociales ($r=-.12$).

Finalmente, la asociación entre las distintas

Tabla 5. Diferencias por tipo de centro asistencial en las puntuaciones HoNOS

	CSM - a (n=164)		HD - b (n=74)		TAC - c (n=22)		HP - d (n=21)		Prueba estadística de contraste de medias		
	M	DE	M	DE	M	DE	M	DE	F	p	Post-Hoc
Probl. conductuales	.84	1.36	.82	1.31	.91	1.85	2.19	2.06	5.606	.001	ad,bd,cd
Deterioro	2.16	1.48	2.39	1.45	1.91	1.60	3.57	1.50	6.201	<.001	ad,bd,cd
Problemas clínicos	3.14	2.22	3.58	2.71	3.27	1.98	2.67	1.62	1.072	.361	
Problemas sociales	4.03	2.86	4.96	2.56	5.36	3.95	9.90	2.26	26.962	<.001	ad,bd,cd
Total	10.18	5.81	11.76	5.47	11.45	5.92	18.33	4.16	13.296	<.001	ad,bd,cd

Nota. CSM=Centro de Salud Mental, HD=Hospital de Día, TAC=Tratamiento Asertivo comunitario, HP=Hospitalización Psiquiátrica, M=Media, DE=Desviación Estándar, Post-Hoc: letras iguales expresan diferencias estadísticamente significativas entre los grupos que representan.

dimensiones del HoNOS han resultado significativas y en sentido positivo para las diferentes combinaciones, salvo en el caso de la asociación entre los problemas clínicos y el deterioro ($r=.10$) que ha resultado estadísticamente no significativa.

Por último, en la Tabla 5 están los resultados del ANOVA para el contraste de diferencias en las puntuaciones del HoNOS atendiendo al dispositivo asistencial donde el participante recibe atención. Para la puntuación total del instrumento los valores más altos se encuentran entre los participantes asistidos en Hospital Psiquiátrico ($M=18.33$), cuya mayor expresión sintomática es estadísticamente significativa respecto a los otros tres tipos de recursos asistenciales (MHospital de Día= 11.76 ; MTratamiento Asertivo Comunitario= 11.45 ; y MCentros de Salud Mental= 10.18), los cuales no difieren entre sí. Un similar patrón de diferencias se encuentra para las tres dimensiones del HoNOS que han resultado estadísticamente significativas (problemas conductuales, deterioro y problemas sociales).

Discusión

La escala HoNOS actualmente es un instrumento ampliamente utilizado en España (Gisbert & Cid-Colom, 2010; Frades et al., 2011; Osta-Samanes et al., 2012; Garay-Arostegui et al., 2014; Escalada et al., 2014, 2015a, 2015b; Ribas et al., 2016) por los clínicos para la valoración de la salud global en los distintos ámbitos asistenciales de la salud mental, a pesar de no contar con la validación española. Sin embargo, a través de los resultados obtenidos en este estudio se ha podido ofrecer evidencias de su adecuación psicométrica como instrumento de valoración clínica.

La consistencia interna de la escala valorada a través del coeficiente alpha de Cronbach es

aceptable ($\alpha=.70$), obteniendo un valor superior a la del estudio original de Wing et al. (1998). Sin embargo, las bajas inter-correlaciones entre los ítems de las dimensiones propuestas originalmente tienen una consistencia interna relativamente pobre, especialmente la dimensión de deterioro (.10), problemas conductuales (.35) y problemas clínicos (.45), salvo la dimensión social que tiene un alfa mayor (.68). Estos valores de alfa de Cronbach han sido encontrados también por otros autores (Preston, 2000; McClelland et al., 2000; Newnham et al., 2009), justificando esta baja intercorrelación entre ítems a la heterogeneidad de la escala (Newnham et al., 2009) y al bajo número de ítems que componen cada dimensión.

Las diferencias por centro asistencial en la gravedad de la enfermedad estarían indicando la capacidad de la escala HoNOS para evaluar este constructo. De hecho, como cabría esperar, teniendo en cuenta el tipo de intervención llevado a cabo por cada uno de los recursos, las medias más altas con diferencia han sido recogidas por los usuarios ingresados en los hospitales psiquiátricos, mientras que las más bajas se recogieron en los centros ambulatorios de Salud mental.

En lo que se refiere a la validez concurrente, la escala EEAG que mide el funcionamiento general de la persona y el EuroQol que evalúa la calidad de vida, correlacionan significativamente de forma negativa con las puntuaciones de las distintas dimensiones del HoNOS, exceptuando el EuroQol con las dimensiones de problemas conductuales y deterioro. Es decir, un peor funcionamiento o calidad de vida (puntuaciones bajas en estas escalas) se asocia a un peor estado de salud o mayor gravedad de la enfermedad en las puntuaciones del HoNOS.

Por el contrario, el HoNOS y la escala de Impresión Clínica Global correlacionan de forma positiva y estadísticamente significativa, o lo que

es lo mismo a mayor mejoría menor gravedad en las distintas dimensiones.

Los resultados de los análisis factoriales confirmatorios pretenden dar respuesta a la controversia existente en relación a la estructura factorial del instrumento, y así poder esclarecer también la unidimensionalidad del mismo. Por ello, se han probado los distintos modelos, resultando ser los modelos bifactoriales los que tienen los índices más ajustados. Entre ellos el modelo bifactorial de cuatro factores correlacionados de Wing et al (1998) y el de cuatro factores correlacionados de Newnham et al. (2009) son los que mejores índices aportan, no habiendo diferencias claras entre ambos modelos. Por tanto, teniendo en cuenta la amplia utilización del instrumento en su estructura original, aceptar la estructura de Wing podría ser la opción más adecuada. Sin embargo, conceptualmente también podría incluirse la propuesta de Newnham et al (2009), ya que los cuatro factores que se proponen (problemas de conducta y antisociales, deterioro físico o psicológico, síntomas y problemas socioeconómicos) han sido agrupados por clínicos y posteriormente probados mediante análisis estadísticos donde se ha encontrado una mayor consistencia interna en tres de los factores, una mejor sensibilidad al cambio y un mejor ajuste en los índices del análisis factorial confirmatorio (Newnham et al., 2009). En este sentido, ambas propuestas cumplen criterios de adecuación matemática y clínica, ya que los dos modelos han sido propuestos por expertos clínicos. No obstante, se considera que la estructura factorial original es la opción que cuenta con mayor trayectoria, ya que ha sido usada extensamente en numerosos estudios tanto clínicos (Jury et al., 2018) como psicométricos (Prabhu & Oakley-Browne, 2008).

Lo que sí permite establecer los resultados del presente estudio es la mayor precisión del modelo bifactorial frente a los modelos simples. Esta estructura bifactorial con cinco factores indica que los ítems distribuyen sus cargas factoriales tanto por dimensiones como en un factor global. Es decir, las puntuaciones que se obtienen a través del HoNOS pueden ser interpretadas tanto por las cuatro dimensiones (problemas conductuales, deterioro, problemas clínicos y problemas sociales), como por una puntuación total de gravedad. Este modelo es acorde al procedimiento

propuesto en el estudio original para interpretación de la herramienta HoNOS (Wing et al., 1998).

Si bien se planificó el estudio en aras a obtener una muestra representativa de la población atendida en la Red de Salud Mental Pública de Bizkaia, los datos obtenidos sólo permiten su generalización, precisamente, a la población de personas con enfermedad mental grave en tratamiento. La enfermedad mental grave por definición es una entidad crónica que, entre las personas en tratamiento, tiende a estabilizarse en sus manifestaciones sintomáticas. Es por esta razón que las valoraciones del HoNOS tienden a mostrar en este estudio bajas puntuaciones, indicando dicha estabilidad dentro de su cronicidad. Es decir, la aplicación de los criterios de inclusión y la selección aleatoria de los participantes han contribuido a que la muestra de estudio se caracterice por ser un grupo de enfermos estables, lo cual produce un sesgo de selección que implica que las características de comportamiento psicométrico del HoNOS sólo sean válidas para este grupo de pacientes. Para una correcta valoración del comportamiento de la herramienta, se hace necesaria la utilización del HoNOS en la valoración, por ejemplo, de personas en situaciones de expresión aguda de síntomas (donde se esperaría altas puntuaciones) o entre personas que estén involucradas en procesos de recuperación (por ejemplo, en programas de incorporación laboral). Asimismo, para verificar su sensibilidad al cambio, también se hacen necesarios estudios longitudinales que permitan valorar la evolución de un paciente desde que se incorpora a un programa de tratamiento.

A este respecto, como ya ha sido realizado en el trabajo llevado a cabo por el grupo de Parabiaghi (Parabiaghi et al., 2005; 2014) en una muestra italiana, sería interesante aplicar las metodologías basadas en la aplicación del Índice de Cambio Fiable propuesto por Jacobson y Truax (1991) que permite conocer si el cambio alcanzado por el individuo se trata de un resultado no solamente estadísticamente significativo sino, también, clínicamente relevante (Iraurgi, 2010).

En conclusión, este estudio aporta datos de la adecuación psicométrica de la versión española de la escala HoNOS, un instrumento ampliamente utilizado en los servicios de salud mental de nuestro entorno, como herramienta clínica para la valoración de la severidad global de los usuarios.

No obstante, se hace necesario buscar más evidencias que permitan valorar a la escala HoNOS como un instrumento de medida de elección en la evaluación de resultados en enfermedad mental grave.

Financiación

El presente trabajo ha sido financiado por el Departamento de Salud del País Vasco en su convocatoria 2013 a Proyectos de Investigación en Salud (Exp.: 2013111088), y a una beca predoctoral del Gobierno Vasco a la tercera autora (PRE_2017_2_0179).

Conflicto de intereses

Los autores declaran no tener ningún conflicto de intereses.

Referencias

- APA - Asociación Americana de Psiquiatría. (2004). *DSM-IV- Manual Diagnóstico y Estadístico de los Trastornos Mentales*. Masson.
- Bebbington, P., Brugha, T., Hill, T., Marsden, L., & Window, S. (1999). Validation of the health of the nation outcome scales. *British Journal of Psychiatry*, 174, 389-394. <https://doi.org/10.1192/bjp.174.5.389>
- Bentler, P.M. (2006). EQS 6 structural equations program manual. Multivariate Software Inc.
- Bobes, J., Portilla, M.P., Bascarán, M.T., Sáiz, P.A., & Bousoño, M. (2002). *Banco de instrumentos psiquiatría clínica*. Psiquiatría Editores.
- Department of Health (1992). Health of the Nation, a strategy for health in England. *White Paper*. HMSO.
- Dogmanas, D. (2014). *Adaptación y validación en población uruguaya de la escala HoNOS de Resultados en Servicios de Salud Mental*. Tesis Doctoral, Universidad de Salamanca.
- Endicott, J., Spitzer, R.L., Fleiss, J.I., & Cohen, J. (1976). The Global Assessment Scale. A procedure for measuring overall severity of psychiatric disturbance. *Archives of General Psychiatry*, 33, 766-771. <https://doi.org/10.1001/archpsyc.1976.01770060086012>
- Escalada, P., & Marín-Fernández, B. (2015a). The nursing diagnosis disturbed thought processes in psychiatric patients: Prevalence and associated characteristics. *International Journal of Nursing Knowledge*, 27, 3, 156-161. <https://doi.org/10.1111/2047-3095>
- Escalada, P., Lazkanotegui, U., Aiolondriz, M., Alameda, M., Flores-del-Redal, C., & Dendarrieta, M. (2015b). Diagnósticos de enfermería, intervenciones y resultados en pacientes institucionalizados con demencia. *Revista Ibero-Americana de Salud y Envejecimiento*, 1, 1, 22-38. [https://doi.org/10.24902/r.riase.2015.1\(1\).22](https://doi.org/10.24902/r.riase.2015.1(1).22)
- Escalada, P., Muñoz-Hermoso, P., González-Fraile, E., Santos, B., González-Vargas, J.A., Fera-Raposo, I. (2014). A retrospective study of nursing diagnoses, outcomes, and interventions for patients with mental disorders. *Applied Nursing Research*, 28, 2, 92-98. <https://doi.org/10.1016/j.apnr.2014.05.006>
- Esteba-Castillo, S., Torrents-Rodas, D., García-Alba, J., Ribas-Vidal, N., & Novell-Alsina, R. (2018). Translation and validation of the Spanish version of the Health of the Nation Outcome Scales for People with Learning Disabilities (HoNOS-LD). *Revista de Psiquiatría y Salud Mental*, 11, 3, 141-150. <https://doi.org/10.1016/j.rpsm.2016.11.002>
- EuroQol Group (1990). EuroQol a new facility for the measurement of health-related quality of life. *Health Policy*, 16, 199-208. [https://doi.org/10.1016/0168-8510\(90\)90421-9](https://doi.org/10.1016/0168-8510(90)90421-9)
- Frades, B., Ruipérez, M.A., & Moro, M. (2011). Estudio descriptivo de los pacientes atendidos en la Unidad de Media Estancia de Salud Mental del Hospital Pare Jofré durante los años 2005-2008. *Rehabilitación Psicosocial*, 8, 1/2, 23-31. <https://doi.org/10.5538/2385-703X.2017.6.5>
- Garay-Arostegui, M., Pousa-Rodríguez, V., & Pérez-Cabeza, L. (2014). La relación entre la percepción subjetiva del funcionamiento cognitivo y el autoestigma con la experiencia de recuperación de las personas con enfermedad mental grave. *Revista de la Asociación Española de Neuropsiquiatría*, 34, 123, 459-475.

- <https://doi.org/10.4321/S0211-57352014000300002>
- Gisbert, C., & Cid-Colom, J. (2010). Evaluación en unidades de rehabilitación hospitalaria: Un balance entre síntomas, funcionalismo, necesidades, expectativas y soporte comunitario. En FEARP (Ed.), *Evaluación en rehabilitación psicosocial*, (pp. 213-231). Federación Española de Asociaciones de Rehabilitación Psicosocial (FEARP).
- Guy, W. (1976). *ECDEU Assessment Manual for Psychopharmacology*. Rockville National Institute Mental Health.
- Guy, W. (2004). Escala de Impresión Clínica Global. En J. Bobes, M. P. García-Portilla, M. T. Bascarán, P. A. Sáiz & M. Bousoño (Eds). *Banco de Instrumentos Básicos para la Práctica de la Psiquiatría Clínica*, 101-102. Psiquiatría Editores.
- Herdman, M., Badia, X., & Berra, S. (2001). EuroQol-5D: Una alternativa sencilla para la medición de la calidad de vida relacionada con la salud en atención primaria. *Atención Primaria*, 28, 6, 425-430.
[https://doi.org/10.1016/S0212-6567\(01\)70406-4](https://doi.org/10.1016/S0212-6567(01)70406-4)
- Iglesias, N., Penas, P., Macía, M., & Iraurgi, I. (2019). *Revisión sistemática internacional de las características psicométricas de la escala HoNOS (Health of the Nation Outcome Scale)*. Comunicación oral. IV Congreso Nacional de Psicología y I International Symposium on Psychological Prevention. Vitoria, España.
- Iraurgi, I. (2010). Evaluación de resultados clínicos (y III): Índices de Cambio Fiable (RCI) en la estimación del cambio clínicamente significativo. *Norte de Salud Mental*, 36, 105-122.
[http://www.ome-aen.org/NORTE/36/11.%20Formación%20continuada%20\(1\)%20Norte36.pdf](http://www.ome-aen.org/NORTE/36/11.%20Formación%20continuada%20(1)%20Norte36.pdf)
- Jacobson, N.S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, 59, 12-19.
<https://doi.org/10.1037/0022-006x.59.1.12>
- Jury, A., Lai, J., Tuason, C., Koning, A., Smith, M., Boyd, L., Swanson, C., Fergusson, D., & Gruar, A. (2018). People who experience seclusion in adult mental health inpatient services: An examination of health of the nation outcome scales scores. *International Journal of Mental Health Nursing*, 28, 1, 199-208. <https://doi.org/10.1111/inm.12521>
- Lorenzo-Seva, U., & Ferrando, P.J. (2019). *Factor 10. Manual of the Program*. Universidad Rovira I Virgili. Recuperado de <http://psico.fcep.urv.es/utilitats/factor>
- Lovaglio, P.G., & Monzani, E. (2012). Health of the nation outcomes scales evaluation in a community setting population. *Quality of Life Research*, 21, 1643-1653.
<https://doi.org/10.1007/s11136-011-0071-9>
- McClelland, R., Trimble, P., Fox, M.L., Stevenson, M.R., & Bell, B. (2000). Validation of an outcome scale for use in adult psychiatric practice. *Quality Health Care*, 9, 98-105.
<https://doi.org/10.1136/ghc.9.2.98>
- McDowell, I. (2006). *Measuring Health: A Guide to Rating Scales and Questionnaires*, Third Edition. Oxford University Press.
- Mokkink, L. B., Terwee, C. B., Patrick, D. L., Alonso, J., Stratford, P. W., Knol, D. L., de Vet, H. C. (2010). The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health related patient-reported outcomes. *Journal of Clinical Epidemiology*, 63, 737-745.
<https://doi.org/10.1016/j.jclinepi.2010.02.006>
- Newnham, E.A., Harwood, K.E., & Page, A.C. (2009). The subscale structure and clinical utility of the Health of the Nation Outcome Scale. *Journal of Mental Health*, 18, 4, 326-334.
<https://doi.org/10.1080/09638230802522486>
- Osta-Samanes, R., García-Tejedor, A., & Muro, C. (2012). Integración social del anciano institucionalizado. *Gerokomos*, 23, 4, 172-176.
<https://doi.org/10.4321/S1134-928X2012000400006>
- Parabiaghi, A., Barbato, A., D'Avanzo, B., Erlicher, A., & Lora, A. (2005). Assessing reliable and clinically significant change on HoNOS: A method for displaying longitudinal data. *Australasian and New Zealand Journal of Psychiatry*, 39, 719-725.
<https://doi.org/10.1080/j.1440-1614.2005.01656.x>

- Parabiaghi, A., Kortrijk, H.E., & Mulder, C.L. (2014). Defining multiple criteria for meaningful outcome in routine outcome measurement using the Health of the Nation Outcome Scales. *Social Psychiatry and Psychiatric Epidemiology*, 49, 2, 291-305. <https://doi.org/10.1007/s00127-013-0750-7>
- Pirkis, J.E., Burgess, P.M., Kirk, P.K., Dodson, S., Coombs, T.J., & Williamson, M.K. (2005). A review of the psychometric properties of the Health of the Nation Outcome Scales (HoNOS) family of measures. *Health and Quality of Life Outcomes*, 3, 76, 1-12. <https://doi.org/10.1186/1477-7525-3-76>
- Prabhu, R., & Oakley-Browne, M. (2008). The use of the Health of the Nation Outcome Scale in an outreach rehabilitation program. *Australasian Psychiatry*, 16, 3, 195-199. <https://doi.org/10.1080/10398560701784813>
- Preston, N.J. (2000). The Health of the Nation Outcome Scales: Validating factorial structure and invariance across two health services. *Australasian and New Zealand Journal of Psychiatry*, 34, 512-519. <https://doi.org/10.1046/j.1440-1614.2000.00726.x>
- Reise, S.P., Bonifay, W.E., & Haviland, M.G. (2013). Scoring and modeling psychological measures in the presence of multidimensionality. *Journal of Personality Assessment*, 95, 2, 129-140. <https://doi.org/10.1080/00223891.2012.725437>
- Ribas, S., Infante, P., & Fernández-González, D. (2016). Rehabilitation programme of Residential Homes: Compare the results in two different models. *International Journal of Integrated Care*, 16, 6, A43, 1-8. <https://doi.org/10.5334/ijic.2986>
- Rodriguez, A., & Reise, S.P. (2016). Evaluating Bifactor Models: Calculating and interpreting statistical indices. *Psychological Methods*, 21, 2, 137-150. <https://doi.org/10.1037/met0000045>
- Schmitter-Edgecombe, M., McAlister, C., & Weakley, A. (2012). Naturalistic assessment of everyday functioning in individuals with mild cognitive impairment: The day-out task. *Neuropsychology*, 5, 631-641. <https://doi.org/10.1037/a0029352>
- Scholtes, V.A., Terwee, C.B., & Poolman, R.W. (2001). What makes a measurement instrument valid and reliable? *Injury*, 42, 3, 236-240. <https://doi.org/10.1016/j.injury.2010.11.042>
- Slade, M. (2002). What outcomes to measure in routine mental health services, and how to assess them: A systematic review. *Australasian and New Zealand Journal of Psychiatry*, 36, 6, 743-753. <https://doi.org/10.1046/j.1440-1614.2002.01099.x>
- Speak, B., & Muncer, S. (2015). The structure and reliability of the Health of the Nation Outcomes Scale. *Australasian Psychiatry*, 23, 1, 66-68. <https://doi.org/10.1177/1039856214563851>
- Speak, B. L., & Muncer, S. J. (2016). Factorial structure of the Health of the Nation Outcome Scales: An ordinal confirmatory factor analysis using a national sample of clinician ratings in England. *International Journal of Mental Health Nursing*, 25, 87-98. <https://doi.org/10.1111/inm.12199>
- Trauer, T. (1999). The subscale structure of the Health of the Nation Outcomes Scales (HoNOS). *Journal of Mental Health*, 8, 5, 199-209. <https://doi.org/10.1080/096382339917196>
- Uriarte, J., Beramendi, V., Medrano, J., Wing, J., Beevor, A., & Curtis, R. (1999). Presentación de la traducción al castellano de la Escala HoNoS (Health of the Nation Outcome Scales). *Psiquiatría Pública*, 11, 4, 93-101.
- Walton, M. K., Powers, J. H., Hobart, J., Patrick, D. L., Marquis, P., Vamvakas, S., Isaac, M., Molsen, E., Cano, S. J., & Burke, L. (2015). Clinical outcome assessments: Conceptual foundation—report of the ISPOR clinical outcomes assessment – emerging good practices for outcomes research task force. *Value Health*, 18, 6, 741-752. <https://doi.org/10.1016/j.jval.2015.08.006>
- Wild, D., Eremenco, S., Mear, I., Martin, M., Houchin, C., Gawlicki, M., Hareendran, A., Wiklund, I., Yee-Chong, L., von-Maltzahn, R., Cohen, L., & Molsen, E. (2009). Multinational trials – recommendations on the translations required, approaches to using the same language in different countries, and the approaches to support pooling the data: the

- ISPOR patient-reported outcomes translation & linguistic validation good research practices task force report. *Value Health*, 12, 430-440. <https://doi.org/10.1111/j.1524-4733.2008.00471.x>
- Wing, J. K., Beevor, A. S., Curtis, R. H., Park, S. B. G., Hadden, S., & Burns, A. (1998). Health of the Nation Outcome Scales (HoNOS). *British Journal of Psychiatry*, 172, 11-18. <https://doi.org/10.1192/bjp.172.1.11>
- Wing, J.K., Lelliott, P., & Beevor, A.S. (2000). Progress on HoNOS. *British Journal of Psychiatry*, 176, 392-395. <https://doi.org/10.1192/bjp.176.4.392>
- Wise, E.A. (2004). Methods for analyzing psychotherapy outcomes: A review of clinical significance, reliable change, and recommendations for future directions. *Journal of Personality Assessment*, 82, 1, 50-59. https://doi.org/10.1207/s15327752jpa8201_10

