

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

⁴ Instituto 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 mejoría 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)	Newnhan (2009)	Lovaglio (2012)	Speak (2015)	Speak (2016)
F1 – Problemas conductuales	It01 – Agresividad	1	1	1	1			4
	It02 – Autoagresiones	3	1	2	1			2
	It03 – Uso de sustancias	1	1	1	1			3
F2 – Deterioro	It04 – Disfunción cognitiva	2	2	1	2	1		1
	It05 – Discapacidad Física	2	2	3	2			1
F3 – Problemas Clínicos	It06 – Depresión	3	3	1	3	1		4
	It07 – Alucinaciones/Delirios	5	3	2	2			2
	It08 – Otros Síntomas	3	3	4	3			2
F4 – Problemas sociales	It09 – Relaciones Sociales	3-4	4	1	1	1	1	3
	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=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	α	h2
1	74.0	13.5	9.3	2.5	.7	.42	.82	195.24	2.14	.32	.69	.22
2	96.1	1.1	1.1	.7	1.1	.10	.53	530.00	6.07	.32	.69	.21
3	76.5	10.0	8.9	3.9	.7	.42	.86	204.76	2.05	.20	.70	.15
4	29.9	35.2	26.7	8.2	.0	1.13	.94	83.19	.34	.27	.69	.16
5	39.1	21.0	25.6	11.4	2.8	1.18	1.15	97.46	.55	.17	.71	.12
6	41.3	22.8	24.6	9.6	1.8	1.08	1.10	101.85	.62	.22	.70	.13
7	43.1	27.8	21.0	7.1	1.1	.95	1.01	106.32	.76	.31	.69	.21
8	43.1	14.6	24.2	15.7	2.5	1.20	1.22	101.67	.47	.38	.68	.27
9	13.9	20.6	38.4	23.8	3.2	1.82	1.05	57.69	-.19	.56	.65	.43
10	28.1	23.1	30.6	16.7	1.4	1.40	1.11	79.29	.15	.65	.63	.52
11	66.5	12.5	11.4	3.9	5.7	.70	1.17	167.14	1.63	.34	.69	.27
12	54.8	16.4	17.4	6.8	4.6	.90	1.19	132.22	1.12	.38	.68	.27
Dimensiones												
Problemas Conductuales						.94	1.49	158.51	1.81	.38	.35	.43
Deterioro						2.31	1.52	65.80	.20	.30	.10	.29
Problemas Clínicos						3.23	2.31	71.52	.72	.37	.45	.42
Problemas Sociales						4.82	3.21	66.19	.69	.54	.68	.67
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 Scheffe.

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 síntoma (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 multivarida, 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	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.

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 HoNO

	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 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 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., Aolondriz, 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., Feria-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., Badía, 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