

Instrumentos de calidad de vida, funcionamiento, síntomas depresivos o ansiosos y secuelas neurocognitivas diligenciados por un sustituto (Proxy).

Revisión sistemática.

Juan David Hernández¹, María Alejandra Spir¹, Kelly Payares¹, Ana María Posada¹, Fabio Alonso Salinas¹, Héctor Iván García¹, Luz H Lugo-Agudelo¹.

1. Grupo de rehabilitación en salud, Universidad de Antioquía, Medellín, Colombia.

Resumen

Antecedentes y objetivo: La evaluación del estado de salud de los pacientes en el transcurso de una condición de salud, es fundamental para proporcionar una medida de referencia que permita cuantificar la variación en su evolución. Sin embargo, en algunos casos no es posible obtener datos directamente del paciente. Por esta razón, el objetivo del presente estudio fue determinar si los instrumentos que no pueden ser aplicados al paciente, pueden ser diligenciados por un sustituto (proxy) como el cuidador, familiar o personal de salud.

Métodos: Se realizó una revisión sistemática de la literatura. Se encontraron 20 estudios y fueron excluidos 35 debido a que no evaluaron el mismo instrumento entre el paciente y el sustituto o no hubo una medida de correlación entre las respuestas de ambos grupos. Los instrumentos que se revisaron en la presente síntesis fueron: Short Form-36 (SF-36), Montreal Cognitive Assessment (MoCA), WHODAS 2.0, Patient Health Questionnaire 9 (PHQ-9), State-Trait Anxiety Inventory (STAI), Disability Rating Scale (DRS).

Resultados: De los 20 estudios incluidos, doce artículos fueron de la escala SF-36, siete artículos del WHODAS 2.0, un artículo de PHQ-9 y no se encontraron artículos relacionados con los instrumentos DRS, STAI y MoCA. Luego, se categorizaron estos 20 artículos por condiciones de salud. Se encontraron diez artículos relacionados con condiciones neurológicas, tres con condición psiquiátrica, tres sobre el adulto mayor, dos con condiciones cardiorrespiratorias y dos con relación a otros temas.

Conclusión: En general, los niveles de concordancia entre las respuestas de los pacientes y los sustitutos fueron buenos, principalmente cuando se evalúa calidad de vida y funcionamiento con los instrumentos SF-36 y WHODAS respectivamente, con un mayor nivel de acuerdo en los dominios más objetivos y observables como la función y el desempeño físicos y menor nivel de acuerdo en los dominios menos objetivos, como los estados emocionales, afectivos y de autopercepción. En los pacientes que no pueden diligenciar los diferentes instrumentos, el uso de un sustituto puede ayudar a evitar la omisión de las respuestas.

Key words

Proxy, Health-Related Quality Of Life, HRQOL, Depressive Disorder, Anxiety Disorders, Neurocognitive Disorders.

Introducción

La evaluación del estado de salud de los pacientes en el transcurso de una condición de salud es fundamental para proporcionar una medida de referencia que permita cuantificar las

variaciones en el tiempo, relacionadas con la progresión de la enfermedad o con las intervenciones clínicas que se realicen (1). Actualmente las medidas del estado de salud complementan las tradicionales de morbilidad y mortalidad. Estas medidas reflejan el modo en que los individuos se sienten y funcionan en su vida diaria y contienen aspectos importantes para los pacientes (2, 3).

En la literatura se han reportado una gran cantidad de instrumentos que evalúan el estado de salud general, la calidad de vida, el funcionamiento, los síntomas depresivos o ansiosos y las secuelas neurocognitivas en diferentes poblaciones. Estos instrumentos son herramientas útiles para los investigadores que buscan desarrollar y probar modelos de atención más efectivos y seguros (3, 4). En su mayoría se diligencian de manera autoinformada o autorreportada (5).

Sin embargo en algunos pacientes su enfermedad o algunas comorbilidades impiden proporcionar información sobre su estado de salud, lo que dificulta obtener datos del paciente. Incluso, en algunas oportunidades no es posible obtenerla, como en pacientes con secuelas cognitivas causadas por un accidente cerebrovascular, esclerosis múltiple, trauma craneoencefálico, enfermedad de Huntington, enfermedad de Alzheimer o en personas de edad avanzada (6) y es aún más difícil en pacientes críticamente enfermos (1,3).

Por lo anterior, se ha planteado la pregunta sobre si es confiable utilizar o no fuentes alternativas de información, como un sustituto o “*proxy*” del paciente para la aplicación de diferentes instrumentos, en los escenarios en los que no pueden ser diligenciados por el paciente (3, 6).

La palabra *proxy* hace referencia a una persona cercana que puede ser un familiar, cuidador o persona a cargo y se han utilizado ampliamente para maximizar la tasa de respuesta de los diferentes instrumentos que miden desenlaces como calidad de vida y funcionamiento (7). La palabra *proxy* no ha sido traducida al español. Dentro de las posibles traducciones que se encontraron para esta investigación estaban ‘acudientes, sustitutos, representantes, intermediarios, cuidadores, familiares, apoderados’. Para esta investigación se decidió utilizar la palabra sustituto para hacer referencia al término *proxy*.

Idealmente cuando se aplica un instrumento o una encuesta sobre el estado de salud y calidad de vida, los propios pacientes son los más indicados para responder. Sin embargo, cuando el paciente no está disponible por estar hospitalizado o institucionalizado o cuando no puede responder por su condición se tiene la opción que lo hagan los sustitutos, que son personas con capacidad de responder a las preguntas. Lo anterior reduce el sesgo de falta de respuesta y se logra que la encuesta o el instrumento aplicado sean representativos de la población estudiada. La población de edad avanzada o con algún tipo de discapacidad ha sido en la que más se han utilizado instrumentos diligenciados por sustitutos (2, 8).

En las principales encuestas de Medicare en Estados Unidos, las respuestas dadas por sustitutos constituyeron entre el 10% y el 30% de todas las respuestas (10). Sin embargo, los resultados notificados por los sustitutos pueden ser sistemáticamente diferentes de los obtenidos directamente de los pacientes (10). El sesgo de respuesta de los sustitutos es la diferencia entre las respuestas de estos, con las de los pacientes. El impacto del sesgo de respuesta de *proxy* en la validez de las estimaciones es una preocupación importante para los

investigadores, cuando las encuestas se realizan entre poblaciones como adultos mayores o personas con discapacidad, sin dejar de reconocer la importancia de obtener esta información de los sustitutos para establecer objetivos de intervención fundamentales en esta población (10). Así mismo, ofrece una solución práctica a la falta de respuesta y a los datos perdidos atribuibles a las limitaciones en la capacidad de los pacientes para responder por sí mismos, en teoría, aquellos que están con una condición más grave y más vulnerables. (8).

Se ha sugerido que el uso de información alternativa, como la suministrada por los sustitutos, sea preferencialmente para evaluar signos objetivos, como la función física y la movilidad, ya que la concordancia de los síntomas percibidos por el paciente, como el estado de ánimo y el funcionamiento emocional, son más débiles por la subjetividad de éstos (5). En general, los sustitutos parecen proporcionar mejores respuestas sobre información más objetiva que subjetiva (3, 4).

Sin embargo, si las respuestas de los representantes son fiables, independientemente del desenlace que se esté evaluando, el sesgo de falta de respuesta debería reducirse. Dado el importante papel que desempeñan los sustitutos para las poblaciones de estudio que tienen dificultades para diligenciar los instrumentos de autoinforme, es importante evaluar el grado de concordancia de las respuestas de los sustitutos con la de los propios pacientes y medir el grado y la dirección de cualquier sesgo que se pueda presentar (4).

El objetivo de este estudio fue evaluar la concordancia, la correlación y la fiabilidad de instrumentos de calidad de vida, funcionamiento, síntomas depresivos, ansiedad, deterioro cognitivo que no pueden ser aplicados al paciente, y pueden ser diligenciados por un sustituto como el cuidador, familiar o personal de salud.

Métodos

El diseño de este estudio fue una revisión sistemática hecha de acuerdo con los criterios PRISMA (Elementos de información preferidos para revisiones sistemáticas y metaanálisis). Se registró un protocolo para esta revisión en PRÓSPERO ‘International prospective register of systematic reviews’, con la identificación: CRD42022318799.

Preguntas de investigación:

¿Cuáles instrumentos que evalúan calidad de vida (SF36), funcionamiento (WHODAS 2.0), DRS, síntomas depresivos y ansiedad (PHQ 9, STAI) o función neurocognitiva (MoCA), que no pueden ser diligenciados o respondidos por el paciente, pueden ser contestados por un sustituto como un cuidador, familiar o personal de salud?

¿Cuáles son las medidas de concordancia, correlación y fiabilidad entre los pacientes y los sustitutos de estos instrumentos?

Estrategia de búsqueda

Se realizó una búsqueda sistemática en PUBMED, BIREME-LILACS, OVID (Cochrane) y Science Direct, Embase con los términos: Short Form-36, Montreal Cognitive Assessment, WHODAS 2.0, Patient Health Questionnaire 9, State-Trait Anxiety Inventory, Disability Rating Scale, *Proxy*, con su respectivo título de tema médico (MeSH) y sinónimos. Las

búsquedas se realizaron en septiembre de 2020 hasta octubre de 2020 y se actualizaron en julio de 2021. Solo se incluyeron los artículos en idioma de español o inglés.

Criterios de inclusión y exclusión

En esta revisión, se incluyeron los estudios que evaluaron pacientes adultos con cualquier condición, a quienes se les haya aplicado cualquiera de las siguientes escalas: SF-36, MoCA, WHODAS 2.0, PHQ-9, STAI y DRS, así mismo fue requisito que la escala también haya sido diligenciada por un sustituto como un cuidador, familiar o personal de salud y que presenten los puntajes o medidas estadísticas para la comparación.

Se excluyeron de esta revisión, los artículos en los que el instrumento diligenciado por el paciente no sea el mismo que el diligenciado por el sustituto para evaluar el desenlace, así como los artículos escritos en un idioma diferente a español o inglés.

Identificación y extracción de los estudios

Después de eliminar los duplicados, se identificaron un total de 271 artículos. Luego se realizó la selección de los artículos por el título y el resumen. Cada artículo fue revisado por dos evaluadores de manera independiente (MAS, JDH). Se encontró conflicto en 45 de los artículos, que fueron resueltos por un tercer evaluador (AMP). Una vez finalizada esta primera selección se obtuvo un total de 55 artículos y dos evaluadores independientes (MAS, JDH) realizaron la revisión de los textos completos de los 55 artículos seleccionados y un tercer evaluador (AMP) resolvió los conflictos. Finalmente, se seleccionaron 20 artículos que cumplen con los criterios de inclusión. La búsqueda se encuentra en el Anexo 1.

La extracción de datos fue realizada por dos investigadores (JDH, MAS) en donde se registraron el autor y año de publicación, título, revista, país, objetivo del estudio, condición de salud de los pacientes del estudio, población incluida, instrumentos utilizados, análisis de datos, resultados, conclusiones y revisión de calidad. Además, se realizó una tabla para los 35 artículos excluidos a texto completo y la razón de su exclusión. La tabla con las características de los estudios incluidos y excluidos se encuentra en el Anexo 2.

Evaluación de la calidad

La evaluación de la calidad de las revisiones fue realizada de forma independiente por dos revisores con la herramienta desarrollada por el Joanna Briggs Institute (JBI) "Checklist for Systematic Reviews and Research Syntheses" (17). La lista de verificación consta de 11 preguntas, cada una de las cuales debe responderse como "sí", "no", "incierto" o "NA" no aplicable. Para cada revisión, se proporcionó una puntuación media de los dos revisores. La evaluación de la calidad se encuentra en la tabla de características de los estudios incluidos (Anexo 2).

Resumen de la evidencia

La síntesis de la evidencia fue realizada por los cinco investigadores (LHL, AMP, KP, MAS, JDH) con experiencia en revisiones sistemáticas. En esta etapa, la información sobre los instrumentos completados por sustitutos se organizó en cuatro categorías:

1. Instrumentos que evaluaban calidad de vida
2. Instrumentos que evaluaban funcionamiento
3. Instrumentos que evaluaban síntomas depresivos o ansiosos
4. Instrumentos que evaluaban el estado neurocognitivo

Instrumentos incluidos en la síntesis

The Short Form 36 Health Survey Questionnaire (SF-36), este es un instrumento genérico utilizado en la evaluación de la calidad de vida relacionada con la salud (CVRS) (Health-Related Quality of Life - HRQL) que incluye ocho dominios: función física (FF) (10 ítems), desempeño físico (DF) (4 ítems), dolor corporal (DC) (2 ítems), salud general (SG) (5 ítems), vitalidad (VT) (4 ítems), función social (FS) (2 ítems) Desempeño emocional (DE) (3 ítems), Salud mental (SM) (5 ítems). Los puntajes obtenidos en cada ítem se suman para generar un puntaje total para cada dominio y estos se transforman en una escala estándar de 0 a 100, un mayor puntaje corresponde a mejor calidad de vida. Además, se pueden calcular dos escalas de resumen: un resumen de componentes físicos (RCF) y otro de componentes mentales (RCM) (11).

El Montreal Cognitive Assessment (MoCA), es una herramienta creada para la detección temprana del deterioro cognitivo leve que evalúa: memoria a corto plazo, habilidades visoespaciales, funciones ejecutivas, atención, concentración y memoria de trabajo, lenguaje y orientación en tiempo y lugar (12).

El World Health Organization Disability Assessment Schedule (WHODAS), es un instrumento genérico de evaluación de la salud y la discapacidad y tiene seis dominios de funcionamiento, incluye: Comprensión y comunicación, capacidad para moverse en su entorno, cuidado personal, relacionarse con otras personas, actividades de la vida diaria y participación en la sociedad. Los puntajes obtenidos en cada ítem se suman para generar un puntaje total para cada dominio y estos se transforman en una escala estándar de 0 a 100, tener mayor puntaje en esta escala se corresponde con tener mayor discapacidad (13).

El Disability Rating Scale (DRS) evalúa cuatro categorías: reactividad y grado de conciencia, capacidad cognitiva para manejar funciones de autocuidado, dependencia física de otros y adaptabilidad psicosocial. La puntuación máxima que un paciente puede obtener en el DRS es 29, lo que se correlaciona con una discapacidad completa. Una persona sin discapacidad puntuará cero (14).

El Patient Health Questionnaire-9 (PHQ-9) es una escala autoaplicada formada por 9 ítems, con un puntaje total de 27 puntos, a mayor puntaje es mayor la gravedad, utilizada para el diagnóstico, seguimiento y medición de la gravedad de los síntomas depresivos (15).

El State-Trait Anxiety Inventory (STAI) mide la ansiedad de los adultos, y diferencia entre "ansiedad de estado" (temporal) y "ansiedad de rasgo" (de larga duración). Contiene dos subescalas: la Escala de Estado de Ansiedad (S-Anxiety) y la Escala de Rasgo de Ansiedad

(T-Anxiety), cada una con 20 ítems, el rango de puntuaciones para cada subprueba es de 20 a 80; la puntuación más alta indica una mayor ansiedad

Síntesis de la evidencia

Las dos medidas que se utilizaron para el análisis de la concordancia, la correlación y la fiabilidad de las respuestas fueron el coeficiente de correlación intraclase (CCI) y las diferencias de medias con sus desviaciones estándar, para estas últimas se calculó el tamaño del efecto con el estadístico Cohen's d, el cual se basa en las diferencias estandarizadas entre las medias, y adicionalmente se estableció para estos resultados su intervalo de confianza (18). Este estadístico se calculó siempre y cuando estuvieran los datos para calcularlo.

El coeficiente de correlación intraclase (CCI) se define como la proporción de variabilidad total debido a la variabilidad de las respuestas de los dos grupos que se están evaluando, en este caso, pacientes y sustitutos. Al ser una proporción los valores del CCI pueden oscilar entre 0 y 1, donde 0 indica ausencia de concordancia y 1 indica concordancia o fiabilidad completa de los resultados presentados (19). Se ha establecido que valores del CCI iguales o inferiores a 0,4 representan una pobre fiabilidad; valores entre 0,41 y 0,70 representan una fiabilidad de moderada a buena, y valores por encima de 0,70 representan una fiabilidad excelente (20).

El tamaño del efecto evaluado con el estadístico "Cohen's d" ayuda a apreciar la magnitud de la diferencia de medias entre las respuestas de los pacientes y sus sustitutos. Siempre que sea posible, se debe proporcionar un intervalo de confianza que acompañe el estadístico para indicar la precisión de la estimación del efecto. Para interpretar los resultados de este estadístico se utilizan como medidas 0.0 – 0.19: Mínimo efecto, 0.20 – 0.49: Efecto pequeño, 0.50 – 0.79: Efecto moderado, > 0.8: Gran efecto (38). En este análisis lo que se buscan son tamaños del efectos mínimos o pequeños lo que indica una menor diferencia entre las respuestas dadas por el paciente al compararlas con las respuestas obtenidas por el sustituto.

Resultados

Descripción de los estudios

Se encontraron 55 estudios y fueron excluidos 35 debido a que no evaluaron el mismo instrumento entre el paciente y el sustituto o no hubo una medida de correlación entre las respuestas de ambos grupos y dos de los estudios son en niños (Ver anexo 2). De los 20 estudios incluidos, doce artículos fueron de la escala SF-36, siete artículos del WHODAS 2.0, un artículo de PHQ-9 (Ver figura 1) y no se encontraron artículos relacionados con los instrumentos STAI, DRS y MoCA (Ver tabla 1). Luego, se categorizaron estos 20 artículos por condiciones de salud. Se encontraron diez artículos relacionados con condiciones neurológicas (22, 23, 21, 24, 30, 31, 32, 29, 28, 35), tres con condición psiquiátrica (26, 33, 34), tres sobre el adulto mayor (25, 3, 4), dos con condiciones cardiorrespiratorias (1, 9), uno en pacientes en UCI (27) y uno con relación a otras condiciones (6) (Ver tabla 1).

Figura 1. Diagrama de flujo de selección de los estudios.

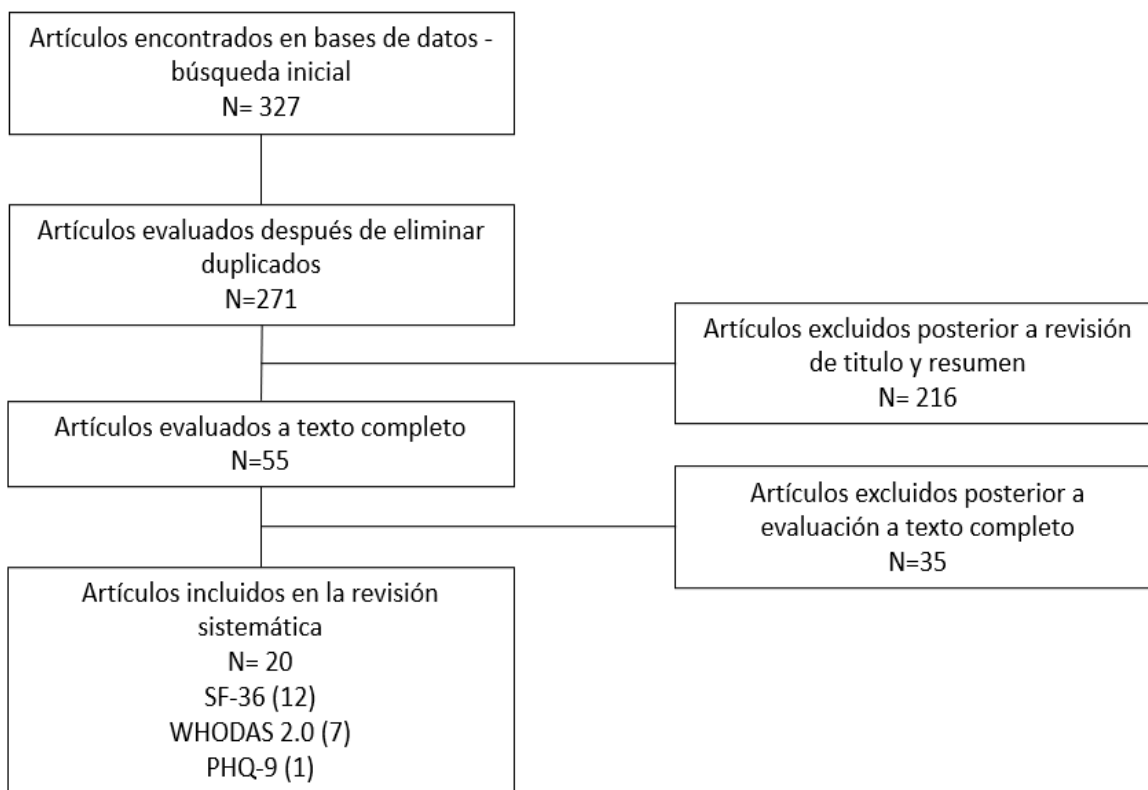


Tabla 1. Artículos categorizados por condiciones de salud.

Condición	Condición específica	Instrumento	N
Condición neurológica (N=10)	Enfermedad de Huntington ²⁸	WHODAS 2.0	1
	Personas con discapacidad ²⁴	SF-36	1
	Enfermedad Cerebrovascular (ECV) ^{29, 35}	PHQ-9 / WHODAS 2.0	2
	Enfermedad de Alzheimer ²²	SF-36	1
	Demencia ²³	SF-36	1
	Esclerosis múltiple ²¹	SF-36	1
	Lesión medular ^{30, 31}	WHODAS 2.0	2
	Trauma encefalocraneano (TEC) ³²	WHODAS 2.0	1
Condición psiquiátrica (N=3)	Esquizofrenia ³⁴	WHODAS 2.0	1
	Trastorno bipolar / Esquizofrenia ²⁶	SF-36	1
	Enfermedad mental ³³	WHODAS 2.0	1
Adultos mayores (N= 3)	Adultos mayores ^{3, 4, 25}	SF-36	3
Condición cardio- respiratoria (N=2)	Condición cardiorrespiratoria ^{1, 9}	SF-36	2
Pacientes en UCI (N=1)	Pacientes en UCI ²⁷	SF-36	1
Otras condiciones (N=1)	Sin enfermedad específica ⁶	SF-36	1

SF-36 Paciente – Sustituto

En los 12 estudios incluidos del SF-36, se evaluaron dos medidas de concordancia principales entre las respuestas del paciente y las del sustituto, la primera medida fue el coeficiente de correlación intraclase (CCI) (Ver tabla 2) y la segunda medida es la evaluación del tamaño del efecto con la diferencia de medias con el “Cohen’s d” (Ver Tabla 3).

En cuanto a las condiciones neurológicas en la esclerosis múltiple todos los dominios presentaron un acuerdo entre moderado a excelente, el dominio que mostró mayor fiabilidad fue la FF con un CCI= 0.91 (IC= 0.87 – 0.95) (21). En los pacientes con Alzheimer el DF fue el de mayor concordancia con un tamaño del efecto $d= 0.07$ (IC= -0.27 - 0.41) en la comparación con familiares y con los cuidadores no familiares; y la FS fue el de menor nivel de acuerdo con $d= 0.56$ (IC= 0.20 - 0.90) (22).

En los pacientes con demencia todos los dominios tienen un pobre nivel de acuerdo en las respuestas proporcionadas por los pacientes al comparar con las de los sustitutos familiar, y cuidador. Aunque no hubo una correlación moderada en ningún dominio, la FF fue el dominio de mayor acuerdo con un CCI= 0.39 y la FS no hubo ningún grado de concordancia entre el paciente y el cuidador (23).

En un grupo de personas con discapacidad, se incluyeron algunos pacientes con esclerosis múltiple, enfermedad de Parkinson y lesión cerebral traumática, hubo diferentes tipos de sustituto (mejor sustituto disponible, familiares, amigos y personal de la salud). La FF, el DC, la SM y la SG obtuvieron una concordancia moderada con todos los tipos de sustitutos. Por el contrario, el DF, el DE y la FS fueron dimensiones que mostraron un pobre nivel de acuerdo (24).

En los estudios que incluían pacientes con condiciones psiquiátricas se encontró que, en la esquizofrenia al comparar las respuestas entre pacientes y sustitutos, en general, todas las dimensiones tuvieron una concordancia de pobre a moderada, el resumen de componentes físicos (RCF) fue el que mostró mayor fiabilidad con un CCI= 0.7 (IC= 0.5–0.8) seguido de los dominios DF, DC y la SM. Cuando se analiza por tamaños de efecto con el Cohen’s d todos los dominios y el resumen de componentes presentan una diferencia pequeña entre las respuestas de los pacientes y los sustitutos, teniendo la menor diferencia el DE con una $d= 0.00$ (IC= -0.5 - 0.5) lo que significa una muy buena concordancia entre ambas respuestas (25).

En los pacientes con trastorno afectivo bipolar (TAB), la concordancia entre pacientes y sustitutos es de moderada a buena en todas las dimensiones, excepto en la dimensión de VT CCI= 0.3. Las dimensiones con mayor acuerdo fueron el DC con CCI=0.8 y la FS con CCI=0.8 y al evaluar tamaños del efecto en las diferencias entre ambos grupos, el dominio de FS igualmente fue el que mostró menor diferencia con una $d= 0.04$ (IC= -0.56 - 0.66), la SM tuvo la mayor concordancia con una $d= 0.01$ (IC= -0.60 - 0.62) (26).

En los pacientes de cirugía cardíaca, la concordancia entre pacientes y sustitutos mostró que la dimensión de DC y la VT tuvieron una pobre concordancia, y éstas se correlacionaron con el tiempo (seguimiento desde antes de la cirugía hasta 6 meses después) (1). Sin embargo, este estudio mostró en el seguimiento a 6 meses que la concordancia fue mayor en los dominios de FF, DF y SM al igual que el RCF (1).

En los pacientes que están en programas de rehabilitación cardíaca fase II cuando se compara las respuestas en la mejoría de CVRS reportada por el paciente con las de su cónyuge, el DE y la SG fueron los dominios de mayor concordancia con una $d= 0.05$ ($-0.37 - 0.48$) en ambos grupos (9).

En una población de adultos mayores, la correlación entre paciente/personal de la salud y la correlación paciente/sustituto indicaron una pobre concordancia en la FS, en los dos escenarios evaluados: servicios de rehabilitación ambulatorios y hospitalarios (3). En ninguna de las dimensiones hubo una concordancia excelente. La dimensión con mejor acuerdo paciente/sustituto fue el FF con $CCI= 0.55$ ($IC= 0.26-0.76$), paciente/profesional de la salud fue de $CCI= 0.45$ ($IC= 0.18 - 0.67$), paciente/sustituto en el entorno de hospital de día $CCI= 0.71$ ($IC= 0.28 - 0.87$) (3).

En un grupo de personas mayores con discapacidad física que evaluaron la concordancia entre el paciente/personal de la salud y el paciente/persona de referencia, al evaluar la CVRS con el SF-36 se encontró que la dimensión de DC tenía la mayor concordancia con un $CCI= 0.69$ seguido del FF con un $CCI=0.6$ (10). Las dimensiones DE y FS tuvieron una concordancia pobre en los dos tipos de sustitutos. Las puntuaciones medias del paciente y del sustituto para los 8 dominios del SF-36, fueron más bajas en el sustituto que las puntuaciones del paciente, con la excepción de la dimensión DF. Las puntuaciones medias de los representantes profesionales en las ocho dimensiones del SF-36 estaban más cerca de las estimaciones de los pacientes (10).

En otro estudio donde se evaluó la correlación en las respuestas del SF-36 en adultos mayores, la mejor concordancia se obtuvo en el FF con $CCI=0.84$ y en el DF cuando se evaluó el tamaño del efecto con una $d= 0.04$ ($IC= -0.45 - 0.53$) mientras que los dominios con más pobre concordancia fueron el DC con $CCI= 0.30$, DE con $CCI=0.31$ y FS con $CCI=0.38$, así como la SM con una $d= 0.52$ ($IC= 0.02 - 1.01$) (4).

Cuando el instrumento SF-36 es diligenciado por un sustituto puede evaluar de forma confiable la calidad de vida del paciente crítico al ingreso en la UCI, el DE fue el de mayor concordancia con una $d= 0.08$ ($IC= -0.18 - 0.34$) y el dominio de menor concordancia fue la SG con $d= 0.43$ ($IC= 0.17 - 0.70$) (27).

En una revisión de Medicare en donde se incluyeron más de 65.000 respuestas de sustitutos de pacientes con diferentes condiciones médicas, todos los dominios del SF-36 tuvieron un tamaño del efecto pequeño al comparar las medias de los pacientes con las respuestas de los sustitutos, lo que significa una excelente concordancia en las respuestas de ambos grupos, con valores de los resúmenes del componente mental $d= 0.02$ ($IC= -0.03 - 0.07$) y del componente físico y 0.11 ($IC= -0.11 - 0.15$). Para los dominios de FF, VT, FS y DE se encontró un tamaño de efecto moderado y para los demás dominios un tamaño de efecto pequeño (6).

Tabla 2. Coeficiente de correlación intraclase (CCI). SF-36 Paciente - Sustituto.

		Coeficiente de correlación intraclase (Intervalo de confianza)									
Población/ autor	Paciente - proxy (N)	(FF)	(DF)	(DC)	(SM)	(DE)	(FS)	(VT)	(SG)	RCM	RCF
Condición Neurológica											
Enfermedad de Alzheimer Novella 2006	Pacientes N= 70/ Familiar proxy N= 63	0.52	0.31	0.25	0.29	0.16	0.07	0.34	0.30		
	Pacientes N= 125/ Cuidador proxy N= 63	0.52	0.34	0.44	0.36	0.22	0.11	0.37	0.42		
Demencia Novella 2001	Pacientes / Familiar proxy N= 125	0.38	0.24	0.27	0.36	0.17	0.10	0.30	0.25		
	Pacientes / Cuidador proxy N= 125	0.39	0.23	0.34	0.35	0.22	0.00	0.36	0.33		
Esclerosis Múltiple Solari, 2001	Paciente / Familiar Proxy (N=243)	0.91* (0.87-0.95)	0.78* (0.72-0.83)	0.84* (0.79-0.89)	0.84* (0.80-0.87)	0.66* (0.58-0.73)	0.73* (0.67-0.78)	0.79* (0.75-0.83)	0.49 (0.41-0.58)		
Personas con Discapacidad Andresen 2001	Paciente/ Mejor proxy disponible (N = 131)	0.58	0.36	0.49	0.48	0.42	0.35	0.51	0.52	0.55	0.54
	Pacientes/ Relativo proxy (N= 78)	0.62	0.32	0.54	0.46	0.38	0.30	0.65	0.55	0.51	0.61
	Pacientes/ Amigos proxy (N= 32)	0.51	0.43	0.63	0.62	0.39	0.53	0.52	0.52	0.63	0.61
	Pacientes/ Cuidador de salud proxy (N= 34)	0.40	0.18	0.46	0.47	0.36	0.3	0.3	0.46	0.49	0.22
Condición psiquiátrica											
Esquizofrenia Kim, 2010	Paciente - proxy Esquizofrenia (N=77)	0.5 (0.3-0.7)	0.6 (0.4-0.8)	0.6 (0.4-0.8)	0.6 (0.4-0.8)	0.4 (0.2-0.6)	0.4 (0.1-0.6)	0.4 (0.0-0.6)	0.5 (0.2-0.7)	0.6 (0.3-0.7)	0.7 (0.5-0.8)
Desorden Bipolar Kim, 2010	Paciente - proxy D. Bipolar (N=50)	0.6 (0.3-0.8)	0.6 (0.3-0.8)	0.8* (0.6-0.9)	0.6 (0.2-0.8)	0.6 (0.3-0.8)	0.8* (0.6-0.9)	0.3 (-0.4-0.6)	0.4 (0.1-0.7)	0.6 (0.2-0.8)	0.7 (0.4-0.8)
Enfermedad cardiovascular											
Enfermedad cardiovascular Elliot 2015	Prequirúrgico - Proxy (N= 96)	0.75 *	0.57	0.38	0.65	0.40	0.52	0.31	0.33	0.51	0.7
	Alta hospitalaria (N= 77)	0.55	0.46	0.19	0.66	0.47	0.25	0.22	0.13	0.64	0.72*
	6 meses posterior al alta (N= 69)	0.81*	0.70	0.28	0.80*	0.75*	0.56	0.32	0.45	0.65	0.80*
Adultos mayores											
Adultos mayores con discapacidad Pierre 1998	Paciente / Profesional de la salud (N=41)	0.38 (0.12-0.58)	0.08 (-0.18-0.32)	0.42 (0.19-0.61)	0.41 (0.18-0.61)	0.13 (-0.13-0.38)	0.01 (-0.24-0.27)	0.60 (0.41-0.74)	0.36 (0.12-0.56)		
	Paciente / Sustituto (N=22)	0.55 (0.26-0.76)	0.40 (0.05-0.68)	0.57 (0.27-0.78)	0.11 (-0.49-0.27)	0.44 (0.10-0.69)	0.19 (-0.15-0.50)	0.11 (-0.25-0.44)	0.58 (0.30-0.78)		
	Paciente / Profesional de la salud (N=38)	0.45 (0.18-0.67)	0.09 (-0.15-0.33)	0.39 (0.15-0.59)	0.41 (0.17-0.61)	0.23 (-0.03-0.46)	0.11 (-0.15-0.36)	0.11 (-0.16-0.37)	0.43 (0.19-0.62)		
	Paciente / Sustituto (N=19)	0.71* (0.28-0.87)	0.03 (-0.42-0.36)	0.21 (-0.20-0.55)	0.52 (0.19-0.75)	0.18 (-0.21-0.52)	0.01 (-0.37-0.38)	0.40 (0.03-0.68)	0.33 (0.01-0.61)		
Adultos mayores con discapacidad Ball 2001	Paciente / Profesional proxy (N = 164)	0.600	0.066	0.690	0.579	0.320	0.333	0.419	0.344		
	Paciente/ lego proxy (N= 164)	0.262	0.102	0.507	0.335	0.204	0.275	0.244	0.308		
Adultos mayores Yip 2001	Paciente - Proxy (N=32)	0.842*	0.502	0.307	0.445	0.318	0.380	0.484	0.688	0.423	0.65

SF-36 dominios: FF: función física; DF: desempeño físico; DC: dolor corporal; SM: salud mental; DE: desempeño emocional; FS: función social; VT: vitalidad; SG: Salud general; RCF: Resumen de componentes físicos; RCM: Resumen de componentes mental. CCI ≤ 0.40 = pobre nivel de acuerdo. CCI 0.41 a 0.70 = Moderado nivel de acuerdo. * CCI ≥ 0.71 = Excelente nivel de acuerdo

Tabla 3. Tamaño del efecto (Cohen´s d) SF-36 pacientes y sustitutos

	Tamaño del efecto: D de Cohen (intervalo de confianza)									
Población, Autor	(FF)	(DF)	(DC)	(SM)	(DE)	(FS)	(VT)	(SG)	RCM	RCF
Condición Neurológica										
Novella 2006 Alzheimer	0.50 (0.15 - 0.84)	0.07 (-0.27 - 0.41)	0.20 (-0.14 - 0.55)	0.23 (-0.11 - 0.58)	0.05 (-0.29 - 0.39)	0.56 (0.20 - 0.90)	0.26 (-0.08 - 0.60)	0.23 (-0.11 - 0.58)		
Condición Psiquiátrica										
Kim 2010 Esquizofrenia	0.28 (-0.22 - 0.78)	0.05 (-0.44 - 0.55)	0.01 (-0.49 - 0.51)	0.01 (-0.49 - 0.50)	0.00 (-0.5 - 0.5)	0.12 (-0.38 - 0.61)	0.07 (-0.43 - 0.57)	0.01 (-0.49 - 0.51)	0.04 (-0.46 - 0.54)	0.05 (-0.44 - 0.55)
Kim 2010 Desorden Bipolar	0.51 (-0.11 - 1.13)	0.09 (-0.52 - 0.70)	0.10 (-0.51 - 0.71)	0.01 (-0.60 - 0.62)	0.20 (-0.41 - 0.81)	0.04 (-0.56 - 0.66)	0.35 (-0.26 - 0.97)	0.22 (-0.39 - 0.83)	0.31 (-0.31 - 0.92)	0.38 (-0.23 - 1.00)
Enfermedad cardiovascular										
Fast 2009 Enfermedad cardiovascular	0.19 (-0.24 - 0.62)	0.12 (-0.30 - 0.55)	0.14 (-0.29 - 0.57)	0.39 (-0.04 - 0.82)	0.05 (-0.38 - 0.48)	0.18 (-0.25 - 0.61)	0.45 (0.02 - 0.89)	0.05 (-0.37 - 0.48)	0.25 (-0.17 - 0.69)	0.12 (-0.31 - 0.54)
Adultos Mayores										
Pierre 1998 Adultos mayores	0.48 (0.13 - 0.84)	0.12 (-0.22 - 0.47)	0.05 (-0.29 - 0.40)	0.00 (-0.35 - 0.35)	0.05 (-0.29 - 0.41)	0.31 (-0.04 - 0.66)	0.22 (-0.13 - 0.58)	0.17 (-0.18 - 0.52)		
Yip 2001 Adultos Mayores	0.19 (-0.30 - 0.68)	0.04 (-0.45 - 0.53)	0.36 (-0.13 - 0.86)	0.52 (0.02 - 1.01)	0.13 (-0.36 - 0.62)	0.27 (-0.22 - 0.77)	0.20 (-0.29 - 0.69)	0.34 (-0.15 - 0.84)	0.36 (-0.13 - 0.86)	0.19 (-0.30 - 0.69)
Otros										
Hofhuis 2003. Pacientes en UCI	0.24 (-0.02 - 0.50)	0.18 (-0.08 - 0.44)	0.18 (-0.08 - 0.44)	0.16 (-0.11 - 0.42)	0.08 (-0.18 - 0.34)	0.26 (0.00 - 0.53)	0.16 (-0.10 - 0.42)	0.43 (0.17 - 0.70)		
Ellis 2003 Sin condición específica	0.17 (0.12 - 0.21)	0.06 (0.01 - 0.10)	0.02 (-0.03 - 0.07)	0.08 (0.04 - 0.13)	0.14 (0.10 - 0.19)	0.15 (0.10 - 0.20)	0.17 (0.12 - 0.22)	0.11 (0.06 - 0.16)	0.02 (-0.03 - 0.07)	0.11 (-0.11 - 0.15)

“Cohen´s d”: 0.0 – 0.19: Mínimo efecto, 0.20 – 0.49: Efecto pequeño, 0.50 – 0.79: Efecto moderado, > 0.8: Gran efecto. SF-36 dominios: FF: función física; DF: desempeño físico; DC: dolor corporal; SM: salud mental; DE: desempeño emocional; FS: función social; VT: vitalidad; SG: Salud general; RCF: Resumen de componentes físicos; RCM: Resumen de componentes.

WHODAS 2.0 Paciente – Sustituto.

De los siete artículos incluidos del WHODAS 2.0, cinco están relacionados con condiciones neurológicas como lesión medular, ECV, TEC y enfermedad de Huntington (28, 29, 30, 31, 32), y dos con condiciones psiquiátricas como esquizofrenia y otros trastornos mentales (33, 34). La evaluación de concordancia en estos estudios se realizó midiendo el tamaño del efecto con la diferencia de medias con el “Cohen´s d” (Ver tabla 4)

En general en las condiciones neurológicas se encontró un tamaño del efecto pequeño en los valores totales del WHODAS 2.0, lo que significa que hay una mayor concordancia entre las evaluaciones del paciente al comparar con las aplicadas por los sustitutos, principalmente en

los pacientes que sufrieron un TEC (32) con una $d=0.02$ (IC= -0.24 - 0.28) y en los pacientes con enfermedad de Huntington (28) con una $d=0.18$ (IC= -0.02 - 0.39). De los dos estudios que evaluaron pacientes con lesión medular no se encontraron resultados similares en cuanto a concordancia, sin embargo, uno de los artículos hacía referencia a pacientes tetraplégicos mientras que en el otro estudio el tipo y nivel de la lesión fueron más variables; En los pacientes con tetraplejía (31) se encontró una buena concordancia entre paciente y sustituto $d=0.11$ (IC= -0.42 - 0.65). De las condiciones neurológicas el único grupo que mostró una gran diferencia en el tamaño del efecto fue el grupo de dolor crónico de columna (30) con una $d=1.92$ (IC= 1.44 - 2.41), y también fue el grupo de todos los estudios incluidos del WHODAS 2.0 con peor concordancia entre las respuestas del paciente y la del sustituto, siendo este último el que percibe mayor alteración del funcionamiento por encima de la percepción del propio paciente (30).

En las condiciones psiquiátricas se incluyeron dos artículos en donde la población del estudio fue pacientes con diagnósticos de esquizofrenia y otros trastornos psicóticos, trastornos del estado de ánimo, trastornos de ansiedad, entre otros (33, 34). En el grupo de pacientes con diagnóstico de esquizofrenia (34) al evaluar el tamaño del efecto entre las respuestas del paciente y sustituto se encontró una buena correlación en las respuestas del WHODAS 2.0 en su puntaje global con una $d=0.01$ (IC= -0.46 - 0.47). Igualmente, en el grupo con diferentes diagnósticos de trastornos mentales con una $d=0.05$ (IC= -0.16 - 0.26) (33)

Tabla 4. Tamaño del efecto (Cohen's d) entre WHODAS 2.0 pacientes y sustitutos

Autor	Población	Cohen's d (Intervalo de confianza)
Condición Neurológica		
Tarvonen-Schröder 2019	Lesión medular	0.67 (0.30 - 1.03)
	Dolor de espalda crónico	1.92 (1.44 - 2.41)
Tarvonen-Schröder 2019	Lesión medular	0.11 (-0.42 - 0.65)
Tarvonen-Schröder 2018	TEC	0.02 (-0.24 - 0.28)
Downing 2014	Huntington	0.18 (-0.02 - 0.39)
Condición psiquiátrica		
Zhou 2020	Desórdenes mentales	0.05 (-0.16 - 0.26)
Pietrini 2021	Esquizofrenia	0.01 (-0.46 - 0.47)

“Cohen's d ”: 0.0 – 0.19: Mínimo efecto, 0.20 – 0.49: Efecto pequeño, 0.50 – 0.79: Efecto moderado, > 0.8: Gran efecto

PHQ-9 Paciente - Sustituto

Se encontró un solo estudio que cumplió con el objetivo de la presente revisión (35). El estudio evaluó la validez y la capacidad de respuesta por parte de los sustitutos en comparación con las respuestas de los pacientes. Este estudio transversal incluyó a 200 pacientes con accidente cerebrovascular. El reporte del PHQ-9 por el paciente encontró un puntaje de 6.0 (+/- 4.9) y el del sustituto fue de 7.0 (+/- 5.4), con un tamaño de efecto de $d=$

0.19 (IC= 0 - 0.39), con una buena concordancia entre las respuestas del paciente y las del sustituto. Finalmente, este estudio concluye que se justifica la utilización de respuestas obtenidas por los sustitutos en pacientes con accidente cerebrovascular con más de 3 meses de evolución (35).

Discusión

Los diferentes instrumentos para la evaluación de la calidad de vida, el funcionamiento, los síntomas depresivos, la ansiedad y el deterioro neurocognitivo están ampliamente validados en la literatura pero aún se presentan algunas dificultades en su aplicación, como en el caso de que no pueden ser diligenciados por el paciente y se pueden utilizar métodos alternativos para obtener esta información como la aplicación del mismo instrumento por un sustituto, cuidador, familiar o personal de salud. Si bien, algunos de los instrumentos están diseñados para ser aplicados tanto a pacientes como a sustitutos, hay otros en los que no esto no está validado.

El SF-36 y el WHODAS 2.0 son los instrumentos que más resultados aportaron para el análisis; por el contrario no se obtuvieron resultados con los instrumentos STAI, DRS y el MoCA. Con el STAI inicialmente se encontraron dos artículos evaluados en niños que tenían como objetivo validar modificaciones de esta escala para ser aplicadas en los padres o sustitutos, sin embargo en general no hubo un buen nivel de acuerdo entre los niños y sus padres con estas escalas modificadas (36). La falta de resultados en la búsqueda del DRS se explica por qué dentro de las características y descripción del instrumento se informa sobre la posibilidad de ser aplicado a sustitutos sin afectar su confiabilidad. Por el contrario, el MoCA al ser un instrumento de evaluación cognitiva no es posible que sea aplicado a un sustituto del paciente y se debe establecer el resultado del instrumento con lo que haya sido posible responder por el paciente (12, 14).

En la evaluación de la calidad de vida se encontró que los dominios que evalúan el componente físico en general y principalmente en condiciones neurológicas son los que muestran mayor nivel de acuerdo y mayor precisión. Por el contrario se encontró un mayor desacuerdo en los dominios del componente mental (DE, SM) y en la FS. Este resultado coincide con lo que se ha descrito en la literatura, y es que hay una mayor concordancia paciente/sustituto en los dominios o elementos que son más visibles y observables como el componente físico. Por el contrario, los dominios y elementos menos observables y más subjetivos como los dominios sociales, ambientales y de autopercepción tienen una menor concordancia en las respuestas paciente/sustituto (6).

Cuando se compara lo encontrado en el SF-36 con otros instrumentos que evalúan calidad de vida como el WHOQOL-BREF en condiciones neurológicas como en el trauma encefalocraneano, se encuentra que con este instrumento el nivel de concordancia sustituto/paciente fue adecuado, teniendo mayor nivel de acuerdo el dominio de funcionamiento físico y menor acuerdo los dominios que evalúan aspectos sociales y de autopercepción. Adicionalmente se describe que la edad del paciente, la gravedad de la lesión y la relación del representante con el paciente pueden afectar el nivel de acuerdo (37).

Los resultados obtenidos con el WHODAS 2.0, una herramienta genérica que mide las actividades y la participación con preguntas más objetivas para el paciente y el sustituto (32),

encontró una muy buena fiabilidad entre las respuestas paciente/sustituto en las condiciones neurológicas y psiquiátricas.

En un estudio en pacientes que sufrieron una ECV se evaluó la respuesta tanto en pacientes como en sustitutos del Rankin modificado (fuerza), índice de Barthel (actividades de la vida diaria), evaluación de Lawton (actividades instrumentales de la vida diaria), Minimal (cognición) y el SIS (Stroke Impact Scale) y se encontró que el sesgo indirecto hacia la sobrevaloración de la gravedad del estado del paciente tendía a aumentar a medida que aumentaba la gravedad del accidente cerebrovascular, pero al evaluar el tamaño del efecto entre las respuestas dadas por el paciente y las dadas por el sustituto fue pequeña (rango, -0,1 a 0,4). con un coeficiente de correlación intraclass que estaba entre 0,50 y 0,83. Además aclaran que el grado de acuerdo fue mejor para los dominios físicos observables. (39)

Con el PHQ-9 solo se encontró un artículo en donde se comparó el instrumento entre pacientes con ECV y sus sustitutos con una buena concordancia en las respuestas de ambos grupos (35), sin embargo, al ser un solo estudio incluido en la evaluación de este instrumento no es posible definir si es aplicable a cualquier población con la posibilidad de reproducir el mismo resultado con la misma confiabilidad.

Esta revisión aporta resultados importantes a clínicos, investigadores y profesionales de la salud en general para evaluar desenlaces de calidad de vida y funcionamiento por medio de las respuestas a los instrumentos por parte de los sustitutos cuando no pueden ser diligenciados por los pacientes, Lo anterior hace que se pueda tener mayor fiabilidad, menos datos perdidos atribuibles a las limitaciones en la capacidad de los pacientes para responder por sí mismos, controlar mejor los sesgos en una investigación y lograr evaluaciones más completas en la práctica clínica de los pacientes que se encuentran en una condición más grave de salud (10).

Limitaciones

La falta de inclusión de otros instrumentos que evalúan los mismos desenlaces evaluados en esta revisión puede limitar la generalización de los resultados a instrumentos de calidad de vida general o de funcionamiento. Podría haber instrumentos específicos con una posibilidad mayor de acuerdo de los desenlaces reportados por el paciente y por los familiares, pero se necesita un mayor número de estudios con estas características.

En la mayoría de los estudios en donde el nivel de concordancia fue evaluado con el coeficiente de correlación intraclass no se incluyó el intervalo de confianza lo que es importante para definir mejor la precisión de los resultados. Además, la calidad metodológica de los estudios primarios incluidos, la falta de los cálculos de los tamaños de muestra y la variabilidad de medidas de correlación de estos afecta la interpretación de los resultados obtenidos.

Implicaciones para la práctica

El uso de fuentes alternativas para obtener información, como un sustituto se convierte en una solución factible a la falta de respuesta y a los datos perdidos atribuibles a las limitaciones en la capacidad de los pacientes para responder por sí mismos en escenarios clínicos.

Utilizar los instrumentos de aplicación en el cuidado de la salud que permiten el uso de sustitutos para responder y diligenciar cuando el paciente no puede hacerlo, facilita y mejora la calidad de los datos en una la investigación controlando mejor los sesgos asociados a los datos incompletos.

Es importante que en los procesos de validación de un instrumento se incluya el componente de evaluación por sustitutos, para las poblaciones con limitaciones para responder la encuesta que se piensa validar y dar unas mejores recomendaciones al final de los procesos de evaluación.

El seguimiento de programas de rehabilitación que utilicen desenlaces de calidad de vida y funcionamiento se benefician de los resultados de esta investigación porque permite conocer mejor cuales instrumentos y que dominios pueden ser utilizados en la evaluación de los pacientes y en su seguimiento, de esta manera se contribuye a mejorar los programas de intervención

Conclusión

En general, los niveles de concordancia entre las respuestas de los pacientes y los sustitutos fueron buenos, principalmente cuando se evalúa calidad de vida y funcionamiento con los instrumentos SF-36 y WHODAS 2.0 respectivamente, con un mayor nivel de acuerdo en los dominios más objetivos y observables como la función y el desempeño físicos y menor nivel de acuerdo en los dominios menos objetivos, como los estados emocionales, afectivos y de autopercepción. En los pacientes que no pueden diligenciar los diferentes instrumentos, el uso de un sustituto (*proxy*) puede ayudar a evitar la omisión de las respuestas y favorecer la toma de decisiones en la práctica clínica conociendo una visión más integral de los efectos de una intervención o de la evolución de una condición de salud determinada

Referencias

1. Elliott D, Lazarus R, Leeder SR. Proxy respondents reliably assessed the quality of life of elective cardiac surgery patients. *J Clin Epidemiol* [Internet]. 2006;59(2):153–9. Disponible en: <http://dx.doi.org/10.1016/j.jclinepi.2005.06.010>
2. Elena M. Andresen; Victoria J. Vahle; Donald Lollar (2001). Proxy reliability: Health-related quality of life (HRQoL) measures for people with disability. 10(7), 609–619. doi:10.1023/a:101318790359
3. Pierre U, Wood-Dauphinee S, Korner-Bitensky N, Gayton D, Hanley J. Proxy use of the Canadian SF-36 in rating health status of the disabled elderly. *J Clin Epidemiol* [Internet]. 1998;51(11):983–90. Disponible en: [http://dx.doi.org/10.1016/s0895-4356\(98\)00090-0](http://dx.doi.org/10.1016/s0895-4356(98)00090-0)
4. Yip JY, Wilber KH, Myrtle RC, Grazman DN. Comparison of older adult subject and proxy responses on the SF-36 health-related quality of life instrument. *Aging Ment Health* [Internet]. 2001;5(2):136–42. Disponible en: <http://dx.doi.org/10.1080/13607860120038357>
5. Rooney AG, McNamara S, Mackinnon M, Fraser M, Rampling R, Carson A, et al. Screening for major depressive disorder in adults with glioma using the PHQ-9: a comparison

of patient versus proxy reports. *J Neurooncol* [Internet]. 2013;113(1):49–55. Disponible en: <http://dx.doi.org/10.1007/s11060-013-1088-4>

6. Ellis BH, Bannister WM, Cox JK, Fowler BM, Shannon ED, Drachman D, et al. Utilization of the propensity score method: an exploratory comparison of proxy-completed to self-completed responses in the Medicare Health Outcomes Survey. *Health Qual Life Outcomes* [Internet]. 2003;1(1):47. Disponible en: <http://dx.doi.org/10.1186/1477-7525-1-47>

7. Davin B, Joutard X, Paraponaris A. “if you were me”: Proxy respondents’ biases in population health surveys. 2019 [cited 2022 Jun 8]; Available from: <https://halshs.archives-ouvertes.fr/halshs-02036434>

8. Elliott MN, Beckett MK, Chong K, Hambarsoomians K, Hays RD. How do proxy responses and proxy-assisted responses differ from what Medicare beneficiaries might have reported about their health care?: How do proxy responses and proxy-assisted responses differ? *Health Serv Res* [Internet]. 2008;43(3):833–48. Available from: <http://dx.doi.org/10.1111/j.1475-6773.2007.00820.x>

9. Fast YJ, Steinke EE, Wright DW. Effects of attending phase II cardiac rehabilitation on patient versus spouse (proxy) quality-of-life perceptions. *J Cardiopulm Rehabil Prev* [Internet]. 2009;29(2):115–20. Available from: <http://dx.doi.org/10.1097/HCR.0b013e31819a023c>

10. Li M, Harris I, Lu ZK. Differences in proxy-reported and patient-reported outcomes: assessing health and functional status among medicare beneficiaries. *BMC Med Res Methodol* [Internet]. 2015;15(1):62. Disponible en: <http://dx.doi.org/10.1186/s12874-015-0053-7>

11. Ware JE Jr, Sherbourne CD. The MOS 36-Item short-form health survey (SF-36): I. Conceptual framework and item selection. *Med Care* [Internet]. 1992;30(6):473–83. Disponible en: <http://dx.doi.org/10.1097/00005650-199206000-00002>

12. MoCA Test [Internet]. MoCA – Cognitive Assessment. 2019 [citado el 18 de abril de 2022]. Disponible en: <https://www.mocatest.org/the-moca-test/>

13. WHO Disability Assessment Schedule (WHODAS 2.0) [Internet]. Who.int. [citado el 18 de abril de 2022]. Disponible en: <https://www.who.int/standards/classifications/international-classification-of-functioning-disability-and-health/who-disability-assessment-schedule>

14. Disability Rating Scale [Internet]. Tbims.org. [citado el 18 de abril de 2022]. Disponible en: <https://www.tbims.org/combi/drs/>

15. Kroenke K, Spitzer RL, Williams JBW. The PHQ-9: Validity of a brief depression severity measure. *J Gen Intern Med* [Internet]. 2001;16(9):606–13. Available from: <http://dx.doi.org/10.1046/j.1525-1497.2001.016009606.x>

16. Julian LJ. Measures of anxiety: State-trait anxiety inventory (STAI), beck anxiety inventory (BAI), and hospital anxiety and depression scale-anxiety (HADS-A). *Arthritis*

Care Res (Hoboken) [Internet]. 2011;63 Suppl 11(S11):S467-72. Disponible en: <http://dx.doi.org/10.1002/acr.20561>

17. Critical appraisal tools [Internet]. Jbi.global. [cited 2022 Jun 8]. Available from: <https://jbi.global/critical-appraisal-tools>

18. Fritz CO, Morris PE, Richler JJ. Effect size estimates: current use, calculations, and interpretation. *J Exp Psychol Gen* [Internet]. 2012;141(1):2–18. Disponible en: <http://dx.doi.org/10.1037/a0024338>

19. Prieto L, Lamarca R, Casado A. Assessment of the reliability of clinical findings: the intraclass correlation coefficient. *Med Clin (Barc)* [Internet]. 1998 [cited 2022 Jun 8];110(4):142–5. Available from: <https://pubmed.ncbi.nlm.nih.gov/9541905/>

20. Hair JF, Celsi M, Celsi MW, Money A, Samouel P, Page M. *The essentials of business research methods*. 3rd ed. London, England: Routledge; 2015..

21. Solari A, Radice D. Health status of people with multiple sclerosis: a community mail survey. *Neurol Sci* [Internet]. 2001;22(4):307–15. Disponible en: <http://dx.doi.org/10.1007/s10072-001-8173-8>

22. Novella JL, Boyer F, Jochum C, Jovenin N, Morrone I, Jolly D, et al. Health status in patients with Alzheimer's disease: an investigation of inter-rater agreement. *Qual Life Res* [Internet]. 2006;15(5):811–9. Disponible en: <http://dx.doi.org/10.1007/s11136-005-5434-7>

23. Novella JL, Jochum C, Ankri J, Morrone I, Jolly D, Blanchard F. Measuring general health status in dementia: practical and methodological issues in using the SF-36. *Aging (Milano)* [Internet]. 2001;13(5):362–9. Disponible en: <http://dx.doi.org/10.1007/bf03351504>

24. Andresen EM, Vahle VJ, Lollar D. Proxy reliability: health-related quality of life (HRQoL) measures for people with disability. *Qual Life Res* [Internet]. 2001;10(7):609–19. Disponible en: <http://dx.doi.org/10.1023/a:1013187903591>

25. Ball AE, Russell EM, Seymour DG, Primrose WR, Garratt AM. Problems in using health survey questionnaires in older patients with physical disabilities. *Gerontology* [Internet]. 2001;47(6):334–40. Disponible en: <http://dx.doi.org/10.1159/000052824>

26. Kim EJ, Song DH, Kim SJ, Park JY, Lee E, Seok JH, et al. Proxy and patients ratings on quality of life in patients with schizophrenia and bipolar disorder in Korea. *Qual Life Res* [Internet]. 2010;19(4):521–9. Disponible en: <http://dx.doi.org/10.1007/s11136-010-9617-5>

27. Hofhuis J, Hautvast JLA, Schrijvers AJP, Bakker J. Quality of life on admission to the intensive care: can we query the relatives? *Intensive Care Med* [Internet]. 2003;29(6):974–9. Disponible en: <http://dx.doi.org/10.1007/s00134-003-1763-6>

28. Downing NR, Kim J-I, Williams JK, Long JD, Mills JA, Paulsen JS, et al. WHODAS 2.0 in prodromal Huntington disease: measures of functioning in neuropsychiatric disease. *Eur J Hum Genet* [Internet]. 2014;22(8):958–63. Disponible en: <http://dx.doi.org/10.1038/ejhg.2013.275>

29. Tarvonen-Schröder S, Hurme S, Laimi K. The World Health Organization Disability Assessment Schedule (WHODAS 2.0) and the WHO Minimal Generic Set of domains of functioning and health versus conventional instruments in subacute stroke. *J Rehabil Med* [Internet]. 2019;51(9):675–82. Disponible en: <http://dx.doi.org/10.2340/16501977-25>
30. Tarvonen-Schröder S, Kaljonen A, Laimi K. Comparing functioning in spinal cord injury and in chronic spinal pain with two ICF-based instruments: WHODAS 2.0 and the WHO minimal generic data set covering functioning and health. *Clin Rehabil* [Internet]. 2019;33(7):1241–51. Disponible en: <http://dx.doi.org/10.1177/0269215519839104>
31. Tarvonen-Schröder S, Kaljonen A, Laimi K. Utility of the World Health Organization Disability Assessment Schedule and the World Health Organization minimal generic set of domains of functioning and health in spinal cord injury. *J Rehabil Med* [Internet]. 2019;51(1):40–6. Disponible en: <http://dx.doi.org/10.2340/16501977-2501>
32. Tarvonen-Schröder S, Tenovuo O, Kaljonen A, Laimi K. Usability of World Health Organization Disability Assessment Schedule in chronic traumatic brain injury. *J Rehabil Med* [Internet]. 2018;50(6):514–8. Disponible en: <http://dx.doi.org/10.2340/16501977-2345>
33. Zhou W, Liu Q, Yu Y, Xiao S, Chen L, Khoshnood K, et al. Proxy reliability of the 12-item world health organization disability assessment schedule II among adult patients with mental disorders. *Qual Life Res* [Internet]. 2020;29(8):2219–29. Disponible en: <http://dx.doi.org/10.1007/s11136-020-02474-w>
34. Pietrini F, Tatini L, Santarelli G, Brugnolo D, Squillace M, Bozza B, et al. Self- and caregiver-perceived disability, subjective well-being, quality of life and psychopathology improvement in long-acting antipsychotic treatments: a 2-year follow-up study. *Int J Psychiatry Clin Pract* [Internet]. 2021;25(3):307–15. Disponible en: <http://dx.doi.org/10.1080/13651501.2021.1912358>
35. Lapin BR, Thompson NR, Schuster A, Honomichl R, Katzan IL. The validity of proxy responses on patient-reported outcome measures: Are proxies a reliable alternative to stroke patients' self-report? *Qual Life Res* [Internet]. 2021;30(6):1735–45. Disponible en: <http://dx.doi.org/10.1007/s11136-021-02758-9>
36. Shain LM, Pao M, Tipton MV, Bedoya SZ, Kang SJ, Horowitz LM, et al. Comparing parent and child self-report measures of the state-Trait Anxiety Inventory in children and adolescents with a chronic health condition. *J Clin Psychol Med Settings* [Internet]. 2020;27(1):173–81. Disponible en: <http://dx.doi.org/10.1007/s10880-019-09631->
37. Hwang H-F, Chen C-Y, Lin M-R. Patient-proxy agreement on the health-related quality of life one year after traumatic brain injury. *Arch Phys Med Rehabil* [Internet]. 2017;98(12):2540–7. Disponible en: <http://dx.doi.org/10.1016/j.apmr.2017.05.013>
38. Téllez A, García CH, Corral-Verdugo V. Effect size, confidence intervals and statistical power in psychological research. *Psychol Russ State Art* [Internet]. 2015;8(3):27–46. Available from:
<http://dx.doi.org/10.11621/pir.2015.0303>

39. Duncan PW, Lai SM, Tyler D, Perera S, Reker DM, Studenski S. Evaluation of proxy responses to the Stroke Impact Scale. *Stroke* [Internet]. 2002;33(11):2593–9. Available from: <http://dx.doi.org/10.1161/01.str.0000034395.06874.3e>