

Liz Paola Benítez Brítez
Facultad de Ciencias de la Salud
Universidad del Norte
Rev Med UN 2015 4(1): 77-126

Validación al español de la Escala de Puntuación de la Discapacidad

RESUMEN

Cada vez es más necesario disponer de instrumentos de medida en el ámbito de salud que se puedan utilizar en la práctica clínica y la investigación. Propósito: Validar al español la Disability Rating Scale, a través de la determinación de su validez de criterio, confiabilidad inter-evaluador, y consistencia interna de sus dominios, para su aplicación en pacientes con Traumatismo Cráneo Encefálico. Metodología: Traducción de la DRS al idioma español, por un traductor y consenso de un grupo de tres consultores, el segundo proceso consistió en la evaluación de 20 pacientes que sufrieron Traumatismo Cráneo Encefálico y fueron atendidos en el Centro de Emergencias Médicas, Hospital Nacional de Itauguá, SENADIS y pacientes particulares del Departamento Central, los cuales fueron evaluados con la DRS traducida al español, la escala Rancho los amigos y la escala FIM aplicadas por 2 evaluadores. Los datos fueron procesados con el programa informático Microsoft Excel 2010, en el cual se hallaron la frecuencia y porcentaje de las variables. Se utilizó el coeficiente de correlación de Pearson y el coeficiente de correlación r^2 , para establecer la validez de criterio entre los resultados de las tres escalas, el Test de Student para la confiabilidad inter-evaluador y el alfa de Cronbach para determinar la consistencia interna. Resultados: Los datos de la DRS con la escala FIM y también con la Escala Rancho los amigos obtuvieron un coeficiente de correlación de Pearson de -0.94 y el coeficiente de correlación r^2 de 0,89. La prueba T fue de 0,1149 y el alfa de Cronbach 0,94. Conclusión: Los resultados han objetivado la validez de criterio, la confiabilidad inter-evaluador, y la consistencia interna de los dominios de la escala, por lo que consideramos que constituye un instrumento válido y confiable para la evaluación y el seguimiento de los pacientes con Traumatismo Cráneo Encefálico en proceso de rehabilitación.

Palabras clave: Traumatismo Cráneo Encefálico, Discapacidad, Validez, Fiabilidad.

INTRODUCCIÓN

En 1948, la Organización Mundial de la Salud definió "salud" como el completo estado de bienestar físico, mental y social, desde entonces se han desarrollado numerosas investigaciones para traducir esta información conceptual en métodos objetivos que mediante cuestionarios u otros instrumentos generen escalas o índices para facilitar la medición de las dimensiones del estado de salud. El desarrollo de una escala de evaluación es un proceso laborioso que puede llevar mucho tiempo antes de conseguir una versión definitiva que llene con las expectativas previstas, es por ello que se deben utilizar cuestionarios que ya han demostrado su utilidad en otros estudios, lo que además permite comparar resultados.

Determinar la validez de una escala a un idioma diferente al que fue creado supone garantizar y comprobar que mide apropiadamente y de modo constante en las mismas circunstancias que es sensible a cualquier tipo de cambios de la situación clínica. El modo de preguntar e incluso el lenguaje utilizado son fuentes de sesgo lo que lleva a pensar que una misma cuestión puede o no ser válida en un idioma u otro.

El objetivo de este trabajo es la validación al español de la escala de puntuación de la discapacidad Disability Rating Scale, proponiendo una puesta al día del proceso de validación de instrumentos en salud y en que consiste, partiendo desde su adaptación cultural al medio donde se quiere utilizar y a continuación volver a medir sus características psicométricas.

FORMULACIÓN DEL PROBLEMA

La mayoría de las escalas de evaluación han sido desarrolladas en países de habla inglesa, lo cual genera inconvenientes cuando van a ser utilizadas en países con lenguaje o culturas diferentes, de modo que con su sola traducción no constituye un instrumento de medición confiable. La pregunta de investigación general sería: ¿Cuál es la validez de la DRS en idioma español? Las preguntas de investigación específicas serían: ¿Cuál es la validez de criterio de la Disability Rating Scale en idioma español? ¿Cuál es la confiabilidad inter-evaluador de la Disability Rating Scale en idioma español? ¿Cuál es la consistencia interna de la Escala DRS en idioma español?

JUSTIFICACIÓN

Vivimos en un entorno cada vez más multicultural y multilingüe en el que las escalas se utilizan como apoyo en la toma de decisiones, es por ello que en los últimos años la validación o modificaciones de escalas ha aumentado en todos los ámbitos evaluativos, esto responde al aumento de números de casos complejos que deben ser enfocados desde varios puntos. Aplicar escalas en pacientes de una cultura diferente de la cultura en la cual fue creada, por parte de profesionales cuya cultura también difiere a la de los creadores de dicho elemento de medición, propone una traducción que exige repetir el proceso de validación o certificación de la escala.

El incremento de este proceso en la actualidad, demuestra la relación de culturas e idiomas en el área de salud en el que las escalas, test y cuestionarios asisten diariamente las evaluaciones, diagnósticos y toma de decisiones. Creemos que la validación de la Escala de Valoración de la Discapacidad en idioma español proporcionará a los profesionales un instrumento de medición confiable, para ser utilizado en la evaluación y tratamiento en pacientes con Traumatismo Craneo Encefálico en proceso de rehabilitación.

OBJETIVOS

Objetivo General

- Validar al español la escala de puntuación de la discapacidad Disability Rating Scale.

Objetivos Específicos

- Establecer la validez de criterio de la Disability Rating Scale en idioma español.
- Determinar la confiabilidad inter-evaluador de la Disability Rating Scale en idioma español.
- Establecer la consistencia interna de la escala DRS en idioma español.

MARCO TEÓRICO

Generalidades

- Déficit en el funcionamiento: sustituye al término "Deficiencia" (utilizada por la CIDDM, desde 1980), es la pérdida o anomalía de una parte del cuerpo o de una función fisiológica o mental.
- Limitación en la Actividad: sustituye al término de "Discapacidad" (tal como se

venía utilizando por la CIDDM), son las dificultades que la persona puede tener en la ejecución de las actividades. La limitación en la actividad puede calificarse en distintos grados, en términos de calidad o cantidad, en la manera de extensión o intensidad en que se esperaría la ejecución de la actividad en una persona sin alteración de la salud.

- Actividad: es el desempeño/realización de una tarea o acción por parte de un individuo.
- Restricción en la Participación: sustituye al término minusvalía, son problemas que una persona puede experimentar en su implicación en actividades vitales.

Escalas de Evaluación

El empleo de las escalas de evaluación determina un incremento en el conocimiento de las repercusiones funcionales que el traumatismo tendría sobre el funcionamiento cerebral global, también es un punto de partida en cuanto al establecimiento de un pronóstico del cuadro, así como un apoyo relevante para la toma de ciertas decisiones diagnósticas y terapéuticas. Las escalas de medición suprimen el factor subjetivo del evaluador y ésta es la base de su utilidad en la investigación clínica y científica.

Aunque las bases para el desarrollo de escalas de evaluación parten inicialmente de observaciones clínicas, el posterior empleo de las mismas y su evaluación, mediante diversos estudios, han servido para constatar sus capacidades y validar la eficacia y eficiencia de muchas de ellas en la actividad asistencial en estos pacientes (Johnston).

Ciertos aspectos específicos del traumatismo craneal (TC), tienen que ser tenidos en cuenta durante la evaluación, en la construcción de herramientas, las dimensiones analizadas, el modo de ejecución y los objetivos específicos. Las alteraciones cognitivas tienen una influencia preponderante sobre el devenir, la minusvalía respecto a la inserción profesional y social, es a menudo pesada; el entorno inmediato del lesionado juega un papel importante para hacer frente a las alteraciones cognitivas, psicológicas y conductuales (Metton 306). Las actuales herramientas de evaluación de las alteraciones neuropsicológicas se sitúan esencialmente en el campo de la discapacidad.

Un problema común al que se enfrentan los profesionales dedicados al tratamiento del paciente con secuelas de TCE es la existencia de una gran variedad de medidas de función y discapacidad, ello dificulta la comparación entre diferentes estudios y datos obtenidos de diferentes centros. Para utilizar con garantía un cuestionario es necesario comprobar que

mide apropiadamente y de modo constante en las mismas circunstancias, que es sensible a los cambios de la situación clínica que en la práctica no presenta dificultades de aplicación. Toda esta comprobación de sus características es lo que se llama proceso de validación

Deficiencias Motoras

Los cuadros clínicos de los TCE son muy heterogéneos y agrupan deficiencias motoras que no tienen especificidad, pueden ser evaluadas con la ayuda de herramientas genéricas.

Escala de Funcionamiento Cognitivo "Rancho Los Amigos"

Esta clasificación establecida por Malkmus, evalúa el comportamiento del paciente con su entorno, pero solo en relación con su estado cognitivo. Realiza una descripción de conductas de donde son deducidos los niveles cognitivos; consta de ocho niveles: sin respuesta; respuesta general; respuesta local; confuso-agitado; confuso-inadecuado; confuso-adequado; automático-adequado; decidido-adequado; permite situar el nivel de recuperación de un paciente con TC, en los momentos posteriores al traumatismo (Salter).

Valoración de la Independencia Funcional

Creada en 1987, se diseñó para abordar los problemas de sensibilidad y exhaustividad y proporcionar un sistema uniforme de medición de la discapacidad con el fin de utilizarlo en el sistema de remuneración médica de los Estados Unidos, evalúa la discapacidad física y cognitiva en cuanto a la carga asistencial, es decir, la puntuación de la FIM representa la carga asistencia de una persona. Consta de 18 apartados que valoran 6 aspectos funcionales (cuidado personal, control de esfínteres, movilidad, locomoción, comunicación y cognición social), estos aspectos se encuadran en dos dominios, físico (13 apartados) y cognitivos (5 apartados).

La FIM es junto con el índice de Barthel, uno de los goldstandards con los que son comparados los nuevos instrumentos de medida de incapacidad, el aspecto negativo es que la Uniform Data System for Medical Rehabilitation (UDSMR), someta el empleo de la FIM, al pago de derechos de utilización anuales, incluso fuera de Estados Unidos.

Escala de Puntuación de la Discapacidad

Esta escala fue desarrollada en 1982 por M. Rappaport, ha sido concebida para obtener información cuantitativa sobre una población de traumatizados craneales moderados o graves, se ocupa de la evaluación de la discapacidad. En la investigación de Hall en 1997 fue probada y desarrollada la DRS en un ámbito de rehabilitación en personas que habrían sufrido un TCE. Se diseñó como seguimiento de los progresos en rehabilitación desde el periodo de coma a través de los diferentes estados de conciencia hasta el retorno del paciente en la comunidad. La DRS a la admisión del paciente está relacionada significativamente con el resultado del mismo al año de la lesión (Salguero).

La popularidad de la DRS deriva de que es breve el tiempo de administración entre 5 y 15 minutos demostrado en un estudio por Hall y cols. 1993, su utilidad se aplica a una gran variedad de configuraciones o niveles de atención, su amplia aceptación su uso en la investigación, y su libre disponibilidad en el dominio público en comparación a otras escalas que precisan el pago de derechos para su uso (Bellon). Se ha recomendado como una medida de resultado útil para ensayos clínicos con individuos con lesión cerebral y actualmente está clasificada como una medida suplementaria dentro del Instituto Nacional de Trastornos Neurológicos y Accidentes Cerebrovasculares de Estados Unidos. Es inusual, ya que tiene la capacidad para describir a personas con discapacidad en una amplia variedad de entornos. En comparación a la FIM ambas escalas son genéricas, pero solo la DRS es específica de pacientes con TCE, por lo que su utilización evalúa aspectos cognitivos deficitarios y característicos del daño cerebral traumático.

Esta escala consta de 8 ítems que están dentro de las siguientes cuatro categorías: reactividad y grado de conciencia, capacidad cognitiva para manejar funciones de autocuidados, dependencia física de otros, y adaptabilidad psicosocial. Los tres primeros ítems de la DRS (apertura de ojos, la capacidad de comunicación, y la respuesta motora) reflejan la función del cuerpo, se encuentran dentro de la categoría de Reactividad y grado de conciencia. Los siguientes tres ítems (alimentación, ir al baño, y aseo personal) se refieren a la actividad se encuentran dentro de la categoría de Habilidad cognitiva para actividades de cuidado personal. Los dos últimos elementos (nivel de funcionamiento y la empleabilidad) reflejan la participación, el primero responde a la habilidad de dependencia de otros y el segundo la adaptación psicosocial.

Cada apartado mide una afección cognitiva y motora mediante una valoración numérica que va de 0 a 3 y 0 a 5 a razón de incrementos de $\frac{1}{2}$ a 1 punto. Cuando los pacientes no se ajustan a las definiciones de puntos enteros para la capacidad cognitiva en relación con los apartados de cuidado personal, la dependencia de los demás y la empleabilidad, se pueden asignar puntos medios. La puntuación total o compuesta se calcula sumando las puntuaciones de los 8 apartados, unas puntuaciones más bajas indican menos discapacidad; la puntuación total se puede usar para asignar al sujeto a una de diez categorías de discapacidad que varían entre la puntuación 0, indica sin ningún nivel de discapacidad; 1 discapacidad leve; de 2

Esta escala consta de 8 ítems que están dentro de las siguientes cuatro categorías: reactividad y grado de conciencia, capacidad cognitiva para manejar funciones de autocuidados, dependencia física de otros, y adaptabilidad psicosocial. Los tres primeros ítems de la DRS (apertura de ojos, la capacidad de comunicación, y la respuesta motora) reflejan la función del cuerpo, se encuentran dentro de la categoría de Reactividad y grado de conciencia. Los siguientes tres ítems (alimentación, ir al baño, y aseo personal) se refieren a la actividad se encuentran dentro de la categoría de Habilidad cognitiva para actividades de cuidado personal. Los dos últimos elementos (nivel de funcionamiento y la empleabilidad) reflejan la participación, el primero responde a la habilidad de dependencia de otros y el segundo la adaptación psicosocial.

a 3 discapacidad parcial; 4 a 6 discapacidad moderada; de 7 a 11 moderadamente grave; de 12 a 16 grave; de 17 a 21 extremadamente grave; de 22 a 24 estado vegetativo; de 25 a 29 estado vegetativo extremo. Sus cimientos son establecer una perspectiva integral y detallada sobre la función y la discapacidad, lo que constituya un ámbito claramente interesante para los profesionales del campo de la rehabilitación en tanto facilita la planificación de la estrategia terapéutica y permite su uso a modo de un valioso instrumento para evaluar el proceso rehabilitador.

Existe un límite teórico entre discapacidad y minusvalía según las definiciones de la OMS no siempre es fácil de establecer, este punto es mencionado ya que dentro de la DRS existe un ítem "Adaptación Psicosocial" que corresponde más a la dificultad de la realización de una tarea que a la limitación de una actividad elemental en la vida cotidiana, y pertenece por lo tanto más al campo de la minusvalía que al de la discapacidad.

Ventajas

- Está constituida por apartados que abarcan todas las dimensiones principales de la ICIDH (deficiencia, discapacidad y minusvalía).
- Posibilita evaluación continua desde la recuperación de la lesión hasta la reintegración social.
- Presenta la capacidad de asignar puntuaciones a una categoría de resultado con relativamente poca pérdida de información, además de proporcionar una visión instantánea del estado de discapacidad general de una persona.
- Presenta mayor sensibilidad a los cambios que la Escala del Coma de Glasgow, es más, las puntuaciones de la última citada pueden obtenerse de la DRS, ya que dentro de su primera categoría evalúa los mismos criterios.

Limitaciones

- En 1992 una investigación llevada a cabo en Brasil ha demostrado que los apartados en cada nivel de puntuación carecen de precisión y las subescalas no identifican con claridad aspectos susceptibles de intervención.
- La DRS evalúa solo una función o un cambio funcional general en vez de específicos.
- Es útil como método para caracterizar la gravedad de una muestra y proporciona un medio para comparación con otros grupos pero no es particularmente sensi-

ble a los efectos de los tratamientos creados para mejorar limitaciones funcionales específicas o la participación social

En este punto cabe mencionar que la FIM es un instrumento más sensible para supervisar los cambios específicos (Hammond)

La Escala de Resultados de Glasgow (ERG), junto con la DRS han sido recomendadas para evaluar la capacidad funcional global del individuo con TCE grave, más pueden también ser utilizados en individuos con TCE con moderada, investigaciones recientes demuestran no existe mucha diferencia.

En un estudio reciente la DRS presenta una fuerte correlación con las 2 versiones de la ERG, mostrándose nítidamente poco sensible para detectar alteraciones postraumáticas cuando se compara con la ERG ampliada, los resultados observados indican que la DRS puede ser de especial valor para demostrar progresos de víctimas luego del trauma, pero es más menos adecuada para la medición de consecuencias pos traumáticas menos graves (Cardozo de Sousa). Las medidas de resultados funcionales más recomendadas como medida de resultados primarios para la lesión cerebral grave, son la Escala de Resultados de Glasgow y la DRS (Guy). La DRS ha sido comparada con la Escala Funcionamiento Cognitivo Rancho Los Amigos pero la primera ha demostrado mayor sensibilidad y fiabilidad.

Validación de Escalas de Medición en Salud

Se han desarrollado numerosas investigaciones para contar con métodos objetivos de medición en el área de rehabilitación que mediante cuestionarios u otros instrumentos generen escalas e índices para facilitar la medición de las dimensiones del estado de salud, su finalidad es conseguir la comparabilidad de información.

Medición

Es el proceso de asignar valores a ciertos eventos de la realidad. La dificultad del proceso radica al menos en dos aspectos: que el valor represente realmente el evento que se quiere medir, y que el evento sea expresado en toda su complejidad.

Medición por medio de escalas: Los instrumentos de medición como las escalas se utilizan para cuantificar atributos, cualidades o propiedades, constructos o conceptos completamente teóricos, que es imposible medir o cuantificar de otra forma (Campo). El proceso de construcción de una escala puede esquematizarse en las siguientes secuencias. Estamos ante algo que creemos que realmente existe y que queremos medir.

- Lo que queremos medir es un fenómeno complejo, que no es observable directa-

- mente y que tiene manifestaciones o maneras de expresarse que no son constantes
- Se descompone lo que queremos medir en grupos de síntomas o manifestaciones que siempre están presentes. Esos grupos se llaman factores o dominios
- Se desarrolla una estrategia para medir los factores o dominios, usualmente mediante la construcción de ítem.

Toda escala debe pasar por un proceso de certificación que se denomina validación de escalas que tiene ciertas características o atributos que respaldan su utilización. La mayoría de estas escalas ha sido desarrollada en países de habla inglesa, lo cual genera una serie de inconvenientes cuando se van a utilizar en países con lenguajes o culturas diferentes, lo que supone que para su uso deberá someterse a un ajuste lingüístico que supone la traducción, que implica que se debe repetir el proceso de certificación del instrumento, es decir, de validación de la escala. Para validar un instrumento de medida hay que comenzar por adaptar culturalmente el cuestionario al medio donde se quiere utilizar y a continuación volver a medir sus características psicométricas.

Adaptación transcultural de un instrumento

El desarrollo de estudios internacionales y la diversidad de culturas en algunos países están aumentando progresivamente. Esto requiere tener accesibles instrumentos de evaluación que estén adaptados culturalmente en los distintos países. La Organización Mundial de la Salud (OMS) recomienda desarrollar y utilizar indicadores de la salud estandarizados, exige la investigación multicéntrica internacional tener instrumentos validados en el ámbito de la salud para poder realizar estudios comparativos (Carvajal). Existen criterios para asegurar la calidad del instrumento: traducción, adaptación, validación de las propiedades psicofísicas y psicométricas y redacción de las normas del instrumento en los distintos idiomas.

¿Por qué validar escalas? Teniendo en cuenta los estándares de selección, construcción y aplicación de medidas de evaluación, existen las siguientes recomendaciones:

- Cuando se hacen cambios sustanciales en el formato del instrumento, modo de aplicación, idioma o contenido, se deberá revalidar la escala para las nuevas condiciones, o tener argumentos que apoyen que no es necesaria o posible una validación adicional.

- Cuando se traduce una escala de un idioma o dialecto a otro, debe establecerse su fiabilidad y validez en los nuevos grupos lingüísticos a los cuales será aplicado.
- Cuando se pretende que las dos versiones de una escala en idiomas distintos sean comparables, hay que aportar pruebas de la comparabilidad de los instrumentos.

Relación Costo-Beneficio

Teniendo en cuenta las recomendaciones citadas, se podría optar por realizar una nueva escala en lugar de iniciar una validación, pero existen argumentos que respaldan la validación:

- Puede resultar más económico y rápido, o sea más eficiente, hacer una validación que desarrollar un nuevo instrumento.
- Al utilizarse instrumentos mundialmente aceptados se abre la posibilidad de efectuar estudios entre diferentes países o entre diferentes culturas.
- Las escalas con amplia aplicación clínica suelen ser instrumentos suficientemente probados.
- El desarrollo de una nueva escala implica disponer de recursos técnicos y humanos altamente calificados y con experiencia en el área.

Validación

Proceso que consiste en adaptar culturalmente un instrumento al medio donde se quiere administrar y comprobar sus características psicométricas como: fiabilidad, validez, sensibilidad y factibilidad. La utilización de escalas de evaluación se basa en la psicofísica y la psicometría. La psicofísica aproxima el proceso de cuantificación de la percepción (trasladar a un sistema numérico fenómenos intangibles como son los síntomas o la discapacidad mediante analogías). La psicometría permite estudiar la adecuación de la escala al fenómeno objeto de la medición y la calidad de la medida

Características Psicométricas de un instrumento

Las dos características métricas esenciales para valorar la precisión de un instrumento son la fiabilidad y la validez.

La fiabilidad: Hace referencia al hecho de medir una variable de manera constante, grado con el que un instrumento mide con precisión sin error. Se ocupa de la proporción de variación en las mediciones que es debida a la diversidad de valores que adopta la variable,

se atribuye a las diferencias verdaderas entre los sujetos. La variación que se debe a un error puede obedecer a 2 tipos de errores:

- Sistemático o sesgo: error que se produce de forma sistemática. Por ejemplo, un evaluador puede puntuar siempre por debajo de los otros evaluadores.
- Aleatorio: error que se produce por factores debido al azar. Por ejemplo, por diferentes circunstancias, un evaluador puede dar a veces puntuaciones a veces superiores y otras veces, puntuaciones inferiores a las correctas, es este tipo de error el que más altera la fiabilidad de un instrumento de medición.

La fiabilidad de un instrumento se valora mediante; la consistencia interna, la fiabilidad test-retest o intraobservador y la fiabilidad interobservador.

- Consistencia interna: Se refiere al grado en que los ítems, puntos o reactivos que hacen parte de una escala se correlacionan entre ellos, la magnitud en que miden el mismo constructo. Una escala consistente garantiza que todos sus componentes o ítems midan un solo constructo que es homogéneo. Cuando nos encontramos con escalas que presentan dentro de sus factores sub-escalas que pretenden medir una dimensión diferente del mismo fenómeno, debe evaluarse la consistencia interna de cada una de ellas.
- Fiabilidad test-retest o intraobservador: Consiste en repetir la prueba dos o más veces a los mismos sujetos con el objeto de correlacionar si las respuestas son similares. Mide la estabilidad de las puntuaciones otorgadas por el mismo evaluador con el mismo método en diferentes momentos.
- Fiabilidad interobservador: El objeto es correlacionar los puntajes obtenidos entre distintos observadores, indicando de esta forma, el porcentaje de acuerdo respecto de la medida observada corrigiendo el factor azar.

La validez: Se refiere a la pertinencia de las inferencias realizadas a partir de las mediciones; en otras palabras, validez significa cuán bien un resultado obtenido refleja al fenómeno bajo estudio, o la capacidad del instrumento para medir la cualidad para la que fue construido. No todo instrumento que sea fiable es válido, un instrumento puede ser fiable porque mide una variable de manera constante, pero inválido si no mide el fenó-

meno que quiere medir. La viabilidad, los mejores instrumentos son inservibles si su aplicación resulta difícil, compleja o costosa. Características como el tiempo empleado en la cumplimentación, la sencillez y la amenidad del formato de interés, la brevedad y la claridad de las preguntas, así como la facilidad de la corrección, el registro la codificación y la interpretación de los resultados son aspectos relacionados con la viabilidad. Esta característica se estudia mediante la realización de un estudio piloto en grupo de individuos y los resultados pueden utilizarse para realizar las mediciones oportunas al instrumento de medición.

La metodología para validar un instrumento es poco conocida por profesionales sanitarios, lo que explica el uso indiscriminado de instrumentos sólo adaptados o validados de manera poco consistente. Validar una versión en idioma diferente al original consiste en adaptar culturalmente el cuestionario al medio donde se quiere utilizar y volver a comprobar que conserva las características adecuadas para medir aquello para lo que fue diseñado en un principio (Argimon). Existen numerosas prácticas que se han tomado como validaciones sin serlo: las más frecuentes son efectuar únicamente traducciones, realizar solamente pruebas de concordancia o correlación con los resultados de la medición de otro instrumento, o practicar tan solo pruebas de concordancia entre diferentes evaluadores.

Proceso de Validación de una Escala

La validación de un instrumento es un proceso continuo y dinámico que va adquiriendo más consistencia cuantas más propiedades psicométricas se hayan medido en distintas culturas, con distintas poblaciones y sujetos. La validación de una escala es un proceso complejo, que comprende varias etapas debe además disponer de herramientas estadísticas para efectuar procedimientos que pueden resultar complejos. En general, el proceso de validación de una escala debe tener las siguientes etapas (Sánchez):

- Selección de la Escala
- Traducción
- Pruebas Preliminares para realizar ajustes (de ítems y utilidad)
- Pruebas de su Validez
- Pruebas de su Confiabilidad
- Determinación de su Utilidad

Selección de la Escala

La escala que se va a validar debe ser la mejor disponible, esto implica efectuar una cuidadosa revisión sistemática de la literatura disponible sobre el tema y efectuar una consulta a expertos en el área sobre la que se efectuará la medición, debe quedar suficientemente fundamentado por qué se seleccionó esa escala para ser validada.

Traducción

Cuando la escala original se encuentra en otro idioma, debe someterse a un proceso de traducción, que debe hacerse con cuidado para evitar distorsiones que generen errores en la medición. El objetivo es conseguir que el instrumento sea equivalente a nivel semántico, conceptual, de contenido, técnico y de criterio en distintas culturas. La equivalencia semántica asegura que el significado de cada ítem sea el mismo en cada cultura después de la traducción. La equivalencia conceptual nos afirma que el instrumento mide el mismo constructo teórico en cada cultura.

El primer paso en el proceso de traducción es conformar un comité de revisión; este comité estará compuesto por personas del grupo encargado de la revisión (como mínimo dos personas), con experiencia en el área de aplicación del instrumento y conocimiento del idioma original de la escala. La cantidad de traductores no es tan importante como la calidad de los mismos, la traducción del idioma original al idioma de aplicación se denomina traducción directa; y la traducción en el otro sentido traducción inversa existen diferentes estrategias para garantizar una correcta traducción (Prieto).

Una de las más prácticas y sencillas recomendaciones es disponer de por lo menos dos traducciones directas efectuadas por diferentes individuos (traductor A y traductor B), si se encuentran diferencias importantes en el significado de algunos de los ítems, debe solicitarse a los traductores que en conjunto, lleguen a un acuerdo sobre los puntos discordantes. Si esto no es posible se entrega la versión del traductor A al B y la del B al A para que realicen una traducción inversa, entonces el comité de revisión evaluará si el significado de los ítems re-traducidos coincide con los de la escala original. Otra alternativa es entregar la versión traducida a un tercer evaluador que hasta ahora no haya participado en el proceso y que se encargará de efectuar la traducción inversa. Con estos insumos el comité de revisión establecerá una versión traducida del instrumento. Uno de los sesgos tradicionalmente encontrados en instrumentos validados o adaptados es la ausencia de equivalencia conceptual de relación al instrumento original.

Prueba preliminar para ajustes

La versión traducida que se definió en la fase previa se aplicará a un grupo de pacientes por parte de dos o tres evaluadores diferentes, tanto los pacientes como los evaluadores deberán tener características similares a las del escenario de aplicación final de la escala. En esta fase se analizan aspectos relacionados con los rubros y con la utilidad de la escala.

En relación con las particularidades del rubro se tienen en cuenta:

- Grado de comprensión de los rubros
- Ambigüedad
- Frecuencia de respuestas
- Restricción del rango de respuesta

En relación con la utilidad de la escala deben evaluarse aspectos como:

- El tiempo de diligenciamiento y aplicación requerida para efectuar la medición.
- Necesidad de entrenamiento para aplicar la escala.
- Características del formato del instrumento
- Facilidad para calificar el puntaje final de la escala.

Pruebas de validez

La validez de un instrumento se refiere a su capacidad para medir aquello para lo que ha sido diseñado, hay diferentes dimensiones de la validez de un instrumento: una dimensión lógica o aparente, una de contenido, una de constructo o concepto y una de criterio.

Validez Aparente: Es una forma de validez de contenido que mide el grado en que los ítems miden lo que se proponen. Es el método más sencillo para medir la validez del instrumento. Para establecer la validez de apariencia se deben conformar dos grupos, uno de sujetos que van a ser medidos con la escala y otro de expertos: ellos analizan la escala y dictaminan si ésta realmente parece medir lo que se propone. Se considera la parte del "sentido común" de la validez de contenido que asegura que los ítems del instrumento sean adecuados (Gould). Consiste en que algunas personas, expertas o no, digan si consideran relevantes los ítems incluidos en un instrumento, cuantas más personas intervengan, mayor validez tendrá esta técnica.

Validez de contenido: La validez de contenido es el grado en que la medición abarca la mayor cantidad de dimensiones del concepto que se quiere estudiar; por tanto se considera que un instrumento es válido por su contenido si se contempla todos los aspectos relacionados con el concepto en estudio, esta dimensión se relacionada con la composición del instrumento y valora si este tiene una muestra representativa (ítem), de los componentes del constructo que pretende medir.

Validez de constructo: Evalúa el grado en el que el instrumento refleja la teoría del fenómeno o del concepto que se quiere medir, garantiza que las mediciones que resulten de las respuestas de la escala, puedan ser consideradas como medición del concepto estudiado. La validación de constructo representa el grado en que una medición se relaciona con otras mediciones de

manera consistente con las hipótesis teóricas que definen el fenómeno que se quiere medir, es una alternativa en caso de ausencia de criterio externo o de referencia. Entre otros procedimientos la validez de constructo se puede verificar mediante estudios de validez convergente-divergente, de análisis factorial y validez discriminante.

- Validez convergente, se correlaciona la medición obtenida a través de un instrumento con otro que mide el mismo constructo. Si los resultados en una u otra medida, correlacionan positivamente o convergen significa, que ambos test miden el mismo constructo.
- Validez concurrente, los resultados de un test se comparan con los de otro que mide un constructo diferente.

En teoría ambos test deben correlaciones negativas o inversas, de lo contrario significa, que el test no es suficientemente específico para medir ese constructo en pacientes quirúrgicos.

Validez de criterio: Es el grado de correlación entre un instrumento y otra medida de la variable a estudio que sirva de criterio o de referencia. Cuando hay una medida de criterio aceptada por investigadores del campo se le llama estándar o regla de oro y los nuevos instrumentos que miden el mismo concepto se comparan con esa medida se habla en este caso de Validez de Criterio Concurrente, la comparación debe de hacerse con un instrumento ya validado, sin embargo cuando no existe una medida estándar la comparación se realiza con otro instrumento que sirva de medida comparable o métodos de apreciación clínica subjetiva. Hay dos formas de validez de criterio: la validez concomitante y la validez predictiva (Prous).

- La validez concomitante: mide el grado de correlación entre dos medidas del mismo concepto al mismo tiempo y en los mismos sujetos. Puede medir la correlación con el instrumento en global o de cada ítem. Esta técnica se utiliza también para seleccionar los mejores ítems de un instrumento y realizar las modificaciones en el instrumento que se está validando.
- La validez predictiva mide el grado de correlación entre un instrumento y una medida posterior del mismo concepto o de otro que está estrechamente relacionado. Mide de qué manera un instrumento predice una evolución o un estado posterior.

La validez de una prueba se evalúa mediante los índices de sensibilidad y especificidad.

- Sensibilidad al cambio. Medir la sensibilidad al cambio es de particular importancia cuando se trata de instrumentos diseñados, no tanto para diagnosticar sino para cuantificar atributos, lo cual nos asegura que la escala es buena para medir una condición a lo largo del tiempo. Los test muy sensibles son aquellos que detectan a la mayoría de los individuos enfermos (pocos falsos negativos). En escalas son las que nos permiten evaluar la respuesta a un tratamiento, el procedimiento más usado para evaluar la sensibilidad al cambio consiste en comparar una puntuación inicial con una puntuación posterior, en un momento en el cual se haya modificado la condición clínica.
- Especificidad: es la proporción de individuos sin la enfermedad que tienen un resultado negativo en la prueba. Las pruebas más específicas son aquellas que van descartando variables en la mayoría de los sujetos sanos.

Pruebas de confiabilidad

Como se mencionó antes, la confiabilidad hace referencia a si la escala funciona de manera similar bajo diferentes condiciones, dependientes del mismo instrumento, del tiempo de aplicación y de la persona que hace la medición. Se puede decir que la confiabilidad es una medición del error que puede generar un instrumento al ser inestable y aplicarse en diferentes condiciones (Alarcón). Debe evaluarse la confiabilidad en los siguientes aspectos.

- Relacionadas con el instrumento, sí los ítems que conforman la escala, tienen unos adecuados niveles de correlación entre ellos, conforman una estructura "aglutinada" que le confiere estabilidad al instrumento.
- Relacionadas con el tiempo de aplicación; se debe medir si la escala, cuando se aplica en diferentes momentos, permaneciendo estable la condición que se mide, mantiene un resultado similar en la medición.

Determinación de la utilidad

Este punto hace referencia a la aplicabilidad del instrumento en el escenario real. Si bien, no depende de la aplicación de procedimientos estadísticos, en la validación de una escala debe describirse el tiempo promedio requerido para aplicar el instrumento, la necesidad de condiciones particulares en las cuales haya que poner al sujeto antes de iniciar el proced-

imiento, el grado de capacitación o calificación profesional que requieren quienes se encargarán de aplicar el instrumento, y la forma, método y tiempo requerido para calificar el puntaje de la escala.

Análisis de pruebas diagnósticas

Cuando se diseña o se realiza la validación de una escala para detectar la presencia o ausencia de un determinado proceso, el instrumento en cuestión se denomina válido si clasifica a los sujetos según presenten o no el proceso con pocos errores, por esta razón es importante determinar el grado de similitud entre los resultados obtenidos de la escala y los obtenidos de un criterio externo de referencia fiable y ampliamente aceptado como medida válida (positivo en presencia del proceso y negativo en ausencia del mismo) del diagnóstico del mismo. Las medidas de evaluación permiten clasificar los sujetos en sanos o enfermos según si el valor de la prueba es superior o inferior al del punto de corte o umbral del instrumento.

MATERIALES Y MÉTODOS

Diseño de la Investigación

- Estudio prospectivo, longitudinal y analítico

Delimitación del Estudio

- Delimitación espacial: La investigación se realizó en la República del Paraguay, con pacientes del Centro de Emergencias Médicas, de la SENADIS, del Hospital Nacional de Itauguá y con pacientes particulares del Departamento Central.
- Delimitación Temporal: La investigación se llevó a cabo en los meses de Noviembre y Diciembre del 2014.

Población y muestra

- Población: Todos los pacientes que sufrieron un Traumatismo craneoencefálico y fueron atendidos en el Centro de Emergencias Médicas, en el Hospital Nacional de Itauguá, en la SENADIS y con pacientes particulares del Departamento Central durante del año 2014.
- Muestra: 20 Pacientes que sufrieron Traumatismo Craneo Encefálico y fueron atendidos en el Centro de Emergencias Médicas, Hospital Nacional de Itauguá, SENADIS y pacientes particulares del Departamento Central durante del año 2014.

Criterios de inclusión

- Pacientes de ambos sexos que sufrieron un Traumatismo Craneo Encefálico

- Pacientes que fueron atendidos en el Centro de Emergencias Médicas, Hospital Nacional de Itauguá o SENADIS durante del año 2014.
- Pacientes que dieron su consentimiento para ser incluidos en el estudio

Criterios de Exclusión

- Pacientes que sufrieron otras afecciones
- Pacientes que no fueron atendidos en el Centro de Emergencias Médicas, Hospital Nacional de Itauguá o SENADIS.
- Pacientes que no dieron su consentimiento para ser incluidos en el estudio

Metodología

Se realizó una selección entre las escala más adecuadas para la medición de la discapacidad, luego de una revisión de la literatura disponible sobre la misma y consultas a terapeutas en el área. El idioma original de la Disability Rating Scale es el inglés para lo cual se realizó su traducción al español con el asesoramiento de un traductor. Para garantizar la fiabilidad de la traducción, se constituyó un Grupo de Revisión, conformado por 3 personas con conocimientos en el área de aplicación del instrumento y conocimiento del idioma original de la escala, quienes evaluaron si el significado de los ítems traducidos coincide con el original, el énfasis del proceso no fue la traducción textual sino la traducción del sentido conceptual que cada ítem persigue.

La escala DRS traducida al español, la escala Rancho los amigos y la escala FIM fueron aplicadas en los meses de noviembre y diciembre del 2014 por 2 evaluadores a 20 pacientes que sufrieron Traumatismo Craneo Encefálico y fueron atendidos en el Centro de Emergencias Médicas, Hospital Nacional de Itauguá, SENADISy con pacientes particulares del Departamento Central.

La validez de criterio fue hallada correlacionando los resultados obtenidos con la escala DRS con los obtenidos con la escala Rancho los amigos y la escala FIM. Para esto se utilizó el coeficiente de correlación de Pearson y el coeficiente de correlación r^2 . La confiabilidad ínter evaluador fue hallada correlacionando los datos obtenidos por cada uno de los 2 evaluadores. Para esto se utilizó el Test de Student. La consistencia se estableció correlacionando los puntajes obtenidos de las dimensiones de la escala entre sí, utilizando para ello el alfa de Cronbach.

Variables

- Sexo
- Edad
- Puntaje en la DRS
- Puntaje en la escala rancho los amigos

- Puntaje en la escala FIM

Análisis Estadísticos

Los datos fueron procesados con el programa informático Microsoft Excel 2010. Las características de la muestra se presentan en frecuencia y porcentaje. Para establecer la validez de criterio se correlacionaron los resultados obtenidos con la escala DRS con los obtenidos con la escala Rancho los amigos y la escala FIM, utilizando el coeficiente de correlación de Pearson y el coeficiente de correlación r^2 . Para establecer la confiabilidad ínter evaluador se utilizó el Test de Student. Para determinar la consistencia interna de la escala se correlacionaron los puntajes de las dimensiones entre sí, utilizando para ello el alfa de Cronbach. A partir de los resultados se realizaron los gráficos correspondientes.

Materiales

- Hojas
- Bolígrafos
- Escala de Medición de la Independencia Funcional (FIM).
- Escala de Puntuación de la Discapacidad (DRS).
- Escala Rancho los Amigos
- Notebook HP Mini
- Paquete informático Microsoft 2010.

Recursos Humanos

- Tutor
- Traductor
- 3 fisioterapeutas que conformaron el equipo de revisión de la traducción al español de la escala DRS
- Personal de Enfermería
- Familiares de los pacientes

Asuntos Éticos

Se explicó a los participantes los objetivos y procedimientos del estudio y se pidió su consentimiento para incluirlos en el mismo. Se incluyeron en la investigación sólo a los pacientes que dieron su consentimiento para el efecto. Los datos personales de los pacientes no serán difundidos ni publicados.

Cronograma

- Agosto: Selección del Tema, elaboración del protocolo.
- Setiembre: Corrección del protocolo, traducción de la escala al idioma español, conformación del grupo de revisión, preparación del marco teórico, consenso de la traducción al español de la DRS por

el grupo de revisión, inicio de la recolección.

- Noviembre: corrección del marco teórico, recolección de datos.
- Diciembre: Elaboración de resultados estadísticos, discusión y conclusión.

RESULTADOS

Características de la muestra

1. Distribución por edad
 - 18 a 29 años: 40%
 - 30 a 39 años: 35%
 - 40 y más años: 25%. Ver gráfico 1.
2. Distribución por sexo
 - Mujeres: 10%
 - Hombres: 90%. Ver gráfico 2.
3. Distribución por gravedad
 - Grave: 95%
 - Moderado: 0%
 - Leve: 5 %. Ver gráfico 3.
4. Tiempo de evolución
 - Hasta 1 año: 70%
 - 1-10 años: 15%
 - Más de 10 años: 15%. Ver gráfico 4.
5. Distribución por ocupación
 - Constructor: 30%
 - Técnico: 25%
 - Estudiante: 15%
 - Agricultor: 15%
 - Mecánico: 10%
 - Abogado: 5%. Ver gráfico 5.
6. Distribución por ciudad
 - Acahay: 20%
 - Concepción: 15%
 - Cordillera: 10%
 - Lambare: 10%
 - Ñemby: 10%
 - Aregua: 5%
 - Fernando de la Mora: 10%
 - San Lorenzo: 10%
 - Capiata: 10%. Ver gráfico 6.

Validez de criterio

7. Correlacionando la DRS con la escala Rancho los amigos se obtuvo un coeficiente de correlación de Pearson de -0.94 y el coeficiente de correlación r^2 de 0.89. Ver gráfico 7.
8. Correlacionando la DRS con la escala FIM se obtuvo un coeficiente de correlación de Pearson de -0.94 y el coeficiente de correlación r^2 de 0,89. Ver gráfico 8.

Confiabilidad inter-evaluador

9. El test de Student entre los datos de los dos evaluadores ha sido de 0,1149.

Consistencia Interna de la DRS

10. En cuanto a la fiabilidad de la consistencia interna se obtuvo mediante el Alfa de Cronbach = 0,9453. Así mismo, el alfa de Cronbach de cada uno de los constructos ha sido mayor a 0.90. Ver tabla 5

DISCUSIÓN

- El sexo predominante masculino y la franja etaria mayoritaria de 18 a 25 años son bastante similares a otros estudios realizados en esta misma Institución (Ayala, López, Quijano, Zarco).
- El hecho que el coeficiente de correlación de Pearson sea de -0,94 y el coeficiente de correlación r^2 de 0,89 significa que la escala DRS tiene una excelente correlación inversa con la escala Rancho los amigos en un 89% de los casos evaluados (Salter).
- El hecho que el coeficiente de correlación de Pearson sea de -0,94 y el coeficiente de correlación r^2 de 0,89 significa que la escala DRS tiene una excelente correlación inversa con la escala FIM en un 89% de los casos evaluados (Salter).
- Asimismo los dos resultados anteriores muestran una excelente correlación entre FIM y Rancho los amigos (Laxe).
- El resultado del test de Student de 0,1149 de muestra que no hay grado de significancia, es decir no hay gran diferencia entre los resultados, lo que hace que los datos sean fiables unos con otros (Salter).
- El Alfa de Cronbach de 0,9453 muestra que los constructos están altamente correlacionados, por lo que también existe una excelente consistencia interna de la DRS en idioma español (Campo).

CONCLUSIÓN

La Disability Rating Scale fue traducida por un traductor y revisado por grupo de fisioterapeutas de experiencia

en evaluación de pacientes con Traumatismo Cráneo Encefálico. Esta versión traducida al español, presenta cualidades psicométricas acordes a los criterios de validación a un idioma diferente al de origen, esto queda demostrado mediante los criterios de confiabilidad inter-evaluador, la excelente relación entre los ítems y los dominios de la misma, también la buena correlación que presenta con otros instrumentos ya certificados.

La validez de criterio se obtuvo correlacionando la DRS con la escala Rancho los Amigos y la escala FIM, dando una excelente correlación por medio del coeficiente de Pearson y una correlación también excelente pero negativa de ambas escalas con la DRS por medio de coeficiente r^2 . Los ítems que la conforman tienen unos adecuados niveles de correlación entre ellos, formando una estructura homogénea, que confiere estabilidad al instrumento, demostrado mediante el alfa de Cronbach. La escala aplicada por diferentes evaluadores, presenta resultados muy similares, esto es lo se mide con la confiabilidad inter-evaluador, por medio de la prueba T. Podemos decir que la DRS, es un instrumento de medición válido en el español, por lo que resulta apropiada para su uso en la práctica asistencial e investigación clínica.

RECOMENDACIONES

La DRS evalúa solo una función o un cambio funcional general en vez de específico, es muy útil como método para caracterizar la gravedad de la muestra y proporciona un medio para comparación con otros grupos, pero no es particularmente sensible a los efectos de los tratamientos diseñados para mejorar limitaciones funcionales específicas, por lo que recomendamos:

- Entrenar a los estudiantes de grado en la utilización de esta escala para el seguimiento de la evolución de pacientes con TCE en proceso de rehabilitación.
- Incentivar a los fisioterapeutas el uso de la escala DRS en idioma español como una herramienta válida y confiable para evaluar y seguir la evolución de pacientes con Traumatismo Cráneo Encefálico.
- En el futuro ampliar la investigación a un grupo más amplio de pacientes.

TRABAJOS CITADOS

- Alarcón M. Muñoz S. "Medición en Salud: Algunas Consideraciones Metodológicas." *Revista Médica de Chile* 2008. Impreso.
- Argimon JM, Jiménez J. "Validación de Cuestionarios." Argimon J, Jiménez J. *Métodos de Investigación Clínica y Epidemiológica*. Elsevier, 2004. Impreso.
- Ayala V. *Correlación entre las Escalas Utilizadas en Pacientes con Traumatismo Cráneo Encefálico [Memoria de Grado]*. Universidad del Norte: Asuncion, 2011. Impreso.
- Bellon K., Wright J., Jamison L., et al. "Disability Rating Scale." *Journals Head Trauma Rehabil* 27.6 (2010): 449-451. Impreso.
- Campo A., Oviedo Heidi. "Propiedades Psicométricas de una Escala: la Consistencia Interna." *Revista de Salud Pública* 2008. Impreso.
- Cardozo de Sousa R. "Comparação entre Instrumentos de Mensuração das Consequências do Trauma Crânio-Encefálico." *Rev Esc Enferm USP* 2006. Impreso.
- Carvajal A. Centeno C. Watson R. et al. "¿Cómo Validar un Instrumento de Medida de la Salud?" *An Sist Sanit Navar* 2011. Impreso.
- Gould J. "Medicine's Core Values: Profession Should not Have to Make Decisions Concerning Rationing." *Br Med* 1994. Impreso.
- Guy L. Ronald L. Harvey L. "Outcome Measures for Clinical Trials Involving Traumatically Brain-Injured Patients: Report of a Conference." (1992). Impreso.
- Hammond FM , Grattan KD , Sasser H , Corrigan JD, et al. "Long-Term Recovery Course After Traumatic Brain Injury: A Comparison of the Functional Independence Measure and Disability Rating Scale." *J Head Trauma Rehabil* 2001. Impreso.
- Johnston MV, Keith RA, Hinderer SR. "Measurements Standards for Multidisciplinary Medical Rehabilitation." *Arch Phys Med Rehabil* 1992. Impreso.
- Laxe S. Capdevila E. Castaño B. *Instrumentos de Medida más Frecuentemente Empleados en la Valoración del Traumatismo Craneoencefálico*. Elsevier-Doyma, 2013. Impreso.
- López F. Sevilla E. Guevara D. et al. "Resultado Funcional al Alta de los Traumatismos Craneoencefálicos Graves Ingresados en una Unidad de Daño Cerebral." *Revista de Neurología* 2004. Impreso.
- Metton G, Béthoux F, Calmels P. *Evaluación y Medición en la Medicina Física y Rehabilitación*. Maddison, 2006. Impreso.
- Prieto AJ. "A Method for Translation of Instruments to Other Languages." *Adult Education Quarterly* 1992. Impreso.
- Prous, M. Salvanés F. Ortells L. *Validación de Cuestionarios*. Elsevier Doyma, 2009. Impreso.
- Quijano M. Cuervo M. "Alteraciones Cognoscitivas Después de un Trauma Craneoencefálico." *International Review of Psychiatry* 2011. Impreso.
- Salgueró P, Ruiz de Vargas C, Chacon J.M. *Traumatismo Craneoencefalicos: Escala de Valoración Para la Medida de Resultados en la Rehabilitación*. Impreso.
- Salter K., Jutai J., Teasell R. *Evaluación de Resultados Después de una Lesión Cerebral Adquirida o Traumática*. Misson, 2006. Impreso.
- Sánchez R. Echeverry J. "Validación de Escalas de Medición en Salud." *Rev Salud Pública* 2004. Impreso.
- Zarco M. Echevarría C. Barrera J. *Valoración del Impacto Del Tratamiento Cognitivo, Dentro del Programa de Rehabilitación en Pacientes con Traumatismo Craneoencefálico*. Elsevier-Doyma, 2013. Impreso.