

Revisión estructurada de las medidas de actividades de la vida diaria en personas mayores

M. José Cabañero-Martínez, Julio Cabrero-García, Miguel Richart-Martínez y Carmen Luz Muñoz-Mendoza

Departamento de Enfermería. Universidad de Alicante. Campus de San Vicente del Raspeig. Alicante. España.

Objetivo: evaluar la adecuación conceptual, la aplicabilidad y las propiedades psicométricas de las medidas de actividades de la vida diaria empleadas con personas mayores españolas.

Material y métodos: se realizaron búsquedas electrónicas en bases de datos nacionales (IME e ISOC) e internacionales (MEDLINE, CINAHL, PsycInfo y EMBASE), así como búsquedas manuales y rastreo de referencias.

Resultados: se incluyeron 34 estudios empíricos relativos a 4 instrumentos: el índice de Barthel, el índice de Katz, la escala de incapacidad física de la Cruz Roja y la escala de actividades de cuidado personal del OARS. Los resultados evidencian, en general, un elevado número de versiones para cada escala, con procesos de adaptación transcultural endebles y sin normas de administración ni de interpretabilidad. La reproducibilidad interobservador es el aspecto de la fiabilidad más examinado. La evidencia sobre la validez predictiva es amplia, pero las evidencias sobre la validez concurrente y la responsividad son casi nulas.

Conclusiones: una versión del índice de Barthel y otra del índice de Katz son las que más evidencias han reunido.

Palabras clave

Anciano. Actividades de la vida diaria. Revisión. Escalas. Estudio teórico.

Structured review of activities of daily living measures in older people

Objective: the present article evaluates the evidence relating to the conceptual adaptation, applicability and psychometric properties of activities of daily living measures in Spanish elderly people.

Material and methods: to obtain original documents, electronic searches were carried out in Spanish (IME and ISOC) and international databases (MEDLINE, PsycInfo, CINAHL and EMBASE). Manual searches and reference searches were also conducted.

Results: 34 articles relating to 4 instruments met the inclusion criteria: the Barthel Index, the Katz Index, the Red Cross Scale of Physical Disability and the Subscale of Personal Care Activities from the OARS. Overall, the results show a large number of versions for each instrument with weak transcultural adaptation processes and without standards for their administration or interpretation. The most frequently evaluated reliability measure was interrater reproducibility. The evidence on predictive validity is extensive, but the evidence on concurrent validity and responsiveness is almost nil.

Conclusions: one version of the Barthel index and another one of the Katz index are the two versions about which the largest quantity of information has been obtained.

Key words

The elderly. Activities of daily living. Review. Scales. Theoretical study.

Este estudio ha sido posible gracias a la financiación de la "Red Temática de Investigación Cooperativa sobre cuidados a personas mayores (RIMARED). Grupo de evaluación de resultados de salud" por el Fondo de Investigaciones Sanitarias, Exp: G03/100, cuyo investigador principal ha sido Julio Cabrero García.

Correspondencia: M.J. Cabañero-Martínez.
Departamento de Enfermería. Universidad de Alicante.
Campus de San Vicente del Raspeig. Apdo. 99. 03080 Alicante. España.
Correo electrónico: Mariajose.cabanero@ua.es

Recibido el 6-11-2007; aceptado el 6-2-2008.

INTRODUCCIÓN

Las actividades de la vida diaria (AVD) son el foco principal de la medida de la salud y la calidad de vida en las personas mayores¹. En la investigación con éstas, la inclusión de medidas de AVD es casi tan frecuente como las de la edad, el sexo y el estado civil². Su medida resulta crítica, ya que son un predictor principal de resultados importantes: mortalidad, utilización de servicios sanitarios e institucionalización, incidencia de enfermedades agudas, caídas y lesiones, y deterioro de la capacidad funcional³. La mayoría de los instrumentos de AVD incluyen ítems de ac-

tividades de autocuidados y de movilidad^{4,5}. A pesar de su importancia, presentan una seria limitación derivada de su carácter tan básico: la discapacidad en AVD es escasamente prevalente en poblaciones comunitarias y presenta un efecto techo en poblaciones clínicas, durante el proceso de recuperación de los pacientes⁶. En consecuencia, la investigación gerontológica amplió el espectro de actividades incluyendo las denominadas actividades instrumentales de la vida diaria (AIVD)⁷. La medida de estas últimas plantea nuevos problemas: las variaciones en las actividades componentes y su dependencia del sexo, la cultura y las preferencias personales⁴.

Clásicamente, la medida del estado de salud en términos de funcionamiento incluye, pues, tres grandes dominios: autocuidados, actividades instrumentales y movilidad⁸. La movilidad interior (cambios de las posiciones del cuerpo, deambular dentro de casa) se adscribe comúnmente al dominio de las AVD y la movilidad exterior, al de las AIVD. El interés de este artículo se centra en la medida de las AVD. El compendio crítico de McDowell y Newell⁹ sobre la medida de la salud con cuestionarios concluyó que la mayoría de las medidas disponibles de AVD se habían desarrollado en un relativo aislamiento de otros métodos y no se fundamentaban en una clara base conceptual ni en la evidencia acumulada. La más reciente revisión de Haywood et al¹⁰, aunque se ha focalizado en medidas de estado de salud en personas mayores, sólo ha incluido las medidas multidimensionales y no las específicas de dominio, como por ejemplo las de AVD/AIVD. En la literatura médica española, Badía et al¹¹ incluyeron algunos instrumentos de AVD en su compendio sobre la medida de la salud en España, y otros autores españoles¹²⁻¹⁴ han abordado la medida de las AVD, pero sin focalizarse en la literatura nacional ni realizar un examen sistemático de la literatura médica.

El objetivo de nuestro trabajo es revisar los instrumentos de AVD empleados en la investigación nacional con personas mayores (de 65 y más años de edad), valorando su adecuación conceptual, su aplicabilidad y sus propiedades psicométricas.

MATERIAL Y MÉTODOS

Estrategia de búsqueda

Para la obtención de documentos originales se han realizado búsquedas electrónicas en las bases de datos nacionales IME e ISOC (1980-2006) y en las bases de datos internacionales MEDLINE, CINAHL, PsycInfo y EMBASE (1989-2006).

Las estrategias de búsqueda fueron diseñadas para obtener estudios originales sobre desarrollo, uso y validación de instrumentos de valoración funcional en personas mayores. Se basaron en la combinación de tres términos: edad, funcionamiento físico y propiedades psicométricas de los instrumentos de medida. En un primer momento se reali-

zaron búsquedas a partir de los descriptores de los tesauros de cada una de las bases de datos que estaban relacionados con los términos de búsqueda. En un segundo momento, con el fin de aumentar la sensibilidad del procedimiento, aun a costa de aumentar el ruido, se generaron, con carácter extensivo, listados de palabras clave, asociados a los términos de búsqueda, para recuperar todos los registros que las contuvieran en título o resumen. En ambas estrategias, los descriptores o las palabras clave, según el caso, referidos a un mismo término se combinaron con el operador *booleano* OR u O y los bloques de descriptores o palabras clave, de cada uno de los tres términos definitorios de la búsqueda, se combinaron con el operador AND o Y. Por último, una vez identificados los instrumentos utilizados en España, se realizaron nuevas búsquedas (IME, ISOC y MEDLINE) con los nombres de éstos y con los nombres de los autores que los habían creado, adaptado o utilizado.

Con las búsquedas electrónicas nunca se localizan todos los documentos indexados sobre el tema de interés y es conveniente completarlas con búsquedas manuales¹⁵. En nuestro caso, se seleccionaron revistas en razón de su notoriedad en la publicación de documentos sobre evaluación de resultados en general y gerontología (1997-2006): *Medicina Clínica, Atención Primaria, Geriatrika, Revista Española de Geriatría y Gerontología y Revista Multidisciplinar de Gerontología* (antes *Revista de Gerontología*).

Las referencias bibliográficas de los artículos que cumplían los criterios de inclusión también se revisaron para obtener artículos adicionales.

Criterios de inclusión

Se incluyeron los artículos originales sobre muestras de personas mayores españolas, cuya edad media fuera superior o igual a 65 años y el rango inferior mayor o igual a 55 años, en los que se valorase la adecuación conceptual, la aplicabilidad o la calidad psicométrica de escalas de AVD.

Criterios de exclusión

Se excluyeron los estudios en los que se administraban instrumentos de AVD pero no aportaban evidencias sobre su rendimiento, en cualquiera de los aspectos indicados en el apartado de extracción de datos.

Dos revisores evaluaron los títulos y los *abstracts* para su inclusión o exclusión en la revisión. Los revisores trabajaron de modo independiente; si existía desacuerdo, un tercero resolvía o recomendaba la lectura completa del artículo.

Extracción de datos

La extracción de la información la llevaron a cabo los mismos investigadores que realizaron la selección de artí-

culos originales, previamente entrenados, de manera independiente, resolviendo los desacuerdos por consenso con un tercero. La información extraída se dividió en dos secciones: información sobre el estudio e información relativa a las escalas focales. La información relativa al estudio incluía las características del estudio y la muestra.

La extracción de la información relativa a las escalas siguió los criterios de calidad definidos para las medidas de resultados de salud basados en el paciente^{9,16,17} e incluía tres bloques de información: a) adecuación conceptual; b) aplicabilidad, y c) propiedades psicométricas (fiabilidad, validez y responsividad).

En la adecuación conceptual se consideró la información relativa a la construcción o adaptación de la escala, el número de dimensiones e ítems, el nivel de medida, el tipo de puntuaciones, el modelo de medida y los efectos suelo y techo.

En la aplicabilidad se recogieron los datos relativos al modo de administración y fuente, demandas de la tarea y requerimientos especiales del informante y administrador. También la información relativa a la interpretabilidad de las puntuaciones de las escalas: datos normativos o significado de las puntuaciones en relación con condiciones clínicas, terapéuticas, eventos vitales socialmente reconocidos, valoraciones subjetivas sobre los cambios mínimamente importantes y predicción de eventos relevantes conocidos¹⁶.

En cuanto a las propiedades psicométricas, para valorar la fiabilidad se registró si los estudios aportaban datos sobre consistencia interna, estabilidad test-retest o equivalencia. Para la consistencia interna se recogieron los valores de los coeficientes alfa de Cronbach, KR-20 o las correlaciones entre dos mitades¹⁸. Para la estabilidad prueba-reprueba y la fiabilidad por equivalencia se anotaron los coeficientes de correlación de Pearson, los coeficientes de correlación intraclase (CCI) y los coeficientes kappa.

Respecto a la validez, se consideró la validez de criterio (concurrente y predictiva) y la validez de constructo. Para evaluar la validez de criterio se extrajeron los valores de sensibilidad y especificidad, en el caso de variables dicotómicas o dicotomizadas, y los estadísticos de correlación, en el caso de medidas continuas u ordenadas, entre la medida y un patrón de oro o criterio clínico. Para valorar la validez de constructo se registró si en los estudios se mostraba la estructura factorial (exploratoria y confirmatoria) de los instrumentos, si se evaluaba la capacidad de discriminación del instrumento entre grupos conocidos o si se aportaban evidencias convergentes y divergentes con variables de su red nomológica u otras, como variables sociodemográficas y clínicas.

También se registraron los datos sobre responsividad: estadísticos de tamaño del efecto o medidas de correlación de cambio entre predictores y criterios clínicos¹⁰.

Además de estas evidencias, si la escala era una versión adaptada de otro idioma, se recogió la información sobre el proceso de adaptación transcultural: traducción simple o retrotraducción y prueba piloto¹⁹.

Por último, se elaboró una tabla sumaria de evidencias de todas las escalas relativa a sus propiedades psicométricas y al proceso de adaptación. En el caso de las propiedades psicométricas, esta valoración siguió el criterio de exhaustividad de las evidencias empleado por McDowell y Newell⁹. Para el proceso de adaptación se distinguieron cuatro categorías ordenadas jerárquicamente: no informado, traducción simple, retrotraducción y estudios piloto.

RESULTADOS

Estrategia de búsqueda

Setenta y nueve artículos cumplieron los criterios de inclusión. Las búsquedas electrónicas aportaron 36 documentos y los manuales, 36, con un solapamiento entre ellas de 12 artículos. Las búsquedas por instrumentos y autores aportaron 9 registros adicionales y a partir de referencias bibliográficas se obtuvieron 10 más. De todos ellos, se excluyeron 45 artículos que no presentaban información bibliográfica de la versión utilizada o que referían un original anglosajón. Puede verse un listado con todos ellos en la siguiente dirección: <http://departamento.enfe.ua.es/profesores/miguel/investigacion.htm>. Por tanto, se han incluido 34 artículos.

Instrumentos seleccionados

Los estudios analizados permitieron identificar 4 escalas de AVD: el índice de Barthel (IB), el índice de Katz (IK), la escala de incapacidad física de la Cruz Roja (EFCR) y la escala de actividades de cuidado personal del OARS (ACP-OARS).

El IB²⁰ está constituido por 10 ítems que valoran la capacidad para realizar ciertas actividades sin ayuda. Evalúa la capacidad de comer, moverse de la silla a la cama y volver, realizar el aseo personal, ir al retrete, bañarse, desplazarse, subir y bajar escaleras, vestirse, control intestinal y control urinario. Su puntuación oscila entre 0 (completamente dependiente) y 100 (completamente independiente) y las categorías de respuesta entre 2 y 4 alternativas, con intervalos de 5 puntos²⁰. En el ámbito internacional existen diferentes versiones con modificaciones en las escalas de puntuación^{21,22} y en el número de ítems²³. En cuanto a las adaptaciones en castellano, sólo Baztán et al¹² indican haber realizado una traducción de la versión de Mahoney y Barthel²⁰.

El IK²⁴ evalúa la dependencia/independencia de las personas en 6 AVD: baño (esponja, ducha o bañera), vestido, uso del retrete, transferencia, continencia y alimentación. En la escala original cada actividad se categoriza en tres niveles de funcionamiento: independencia, dependencia parcial y dependencia total. Niveles que posteriormente se clasifican en dos: dependencia o independencia²⁵. Sobre la

Tabla 1. Resumen de los principales datos sobre aplicabilidad y adaptación transcultural de las diferentes versiones en castellano de las escalas de AVD

Instrumento	Contexto	Administración	Fuente	Proceso de adaptación
Índice de Barthel (Mahoney y Barthel, 1965) ²⁰				
Baztán et al, 1993 ^{12,32-45}	1-2-3-6	Entrevista Autoadministrado Desempeño	Paciente Proxy* Ambas	Traducción
Alarcón-Alarcón et al, 1993 ⁴⁶⁻⁴⁸	1	Entrevista	Ambas	No informado
Fontana-Fortuny et al, 2002 ⁴⁹	3	Entrevista	Paciente	No informado
Índice de Barthel (Shah et al, 1989) ²²				
Cid y Damián, 1997 ^{13,50-52}	6	No informado	No informado	No informado
Índice de Barthel (Granger et al, 1979) ²³				
Grup de Treball, 1997 ^{53,54}	6	No informado	No informado	Traducción al catalán
Índice de Katz (Katz et al, 1963) ²⁴				
Álvarez-Solar et al, 1992 ^{26,55,56}	4-5	Entrevista	Paciente	Traducción
Cruz-Jentoft, 1991 ^{57,58}	4	Entrevista	Ambas	No informado
Cruz-Jentoft, 1991 ^{14,59,60}	5-6	Desempeño	Paciente	Traducción
Fontana-Fortuny et al, 2002 ⁴⁹	3	Entrevista	Paciente	No informado
González-Montalvo et al, 1991 ^{27,61}	1-2-6-5	Entrevista	Paciente	Traducción
Escala de incapacidad física de la Cruz Roja				
Guillén-Llera, 1972 ^{28,29,35,47,48}	1-2-5-6	Entrevista	Paciente Ambas	No procede
Guillén-Llera, 1986 ^{62,27,63}	1-2-5	Entrevista	Ambas	No procede
Cruz-Jentoft, 1991 ^{57,58}	4	Entrevista	Ambas	No procede
Escala de actividades de cuidado personal del OARS, Fillenbaum, 1988 ³⁰				
Grau-Fibla et al, 1996 ^{31,64,65}	3	Entrevista	Ambas	Traducción, retrotraducción y estudio piloto

Contextos: 1 (atención especializada con ingreso [agudos]); 2 (atención especializada ambulatoria-hospital de día); 3 (comunidad-centros de día); 4 (atención primaria); 5 (atención a domicilio); 6 (instituciones de media [unidades de convalecencia y rehabilitación] y larga estancia).

*Proxy: allegado.

base de esta última categorización, las personas son incluidas en uno de los 8 niveles de dependencia del índice, que oscilan entre A (independiente para todas las funciones) y G (dependiente en todas las funciones), existiendo un nivel denominado O (dependiente en al menos dos funciones pero no clasificable como C, D, E o F)²⁴. Entre las versiones españolas, Álvarez-Solar et al²⁶ presentan una traducción simple del índice con modificaciones importantes sobre el original. El observador puntúa cada actividad en una escala de 2 niveles (dependiente/independiente) y no incluye definiciones de las funciones a evaluar. Los estudios de González-Montalvo et al²⁷ y Cruz-Jentoft¹⁴ también presentan traducciones simples, pero en estos casos, los autores mantienen las tres opciones de la escala original.

La EFCR²⁸ es un índice global de incapacidad funcional. Evalúa la incapacidad física según una escala de 6 grados, que oscilan entre 0 (se vale totalmente por sí mismo, anda con normalidad) y 5 (inmovilizado en cama o sillón, incontinencia total, necesita cuidados continuos de enfermería). Es un instrumento de elaboración española^{28,29}.

La ACP-OARS³⁰ constituye uno de los dos factores de la escala de capacidad de autocuidado del OARS Multidimensional Functional Assessment Questionnaire³⁰. Evalúa 7 actividades: capacidad para comer, vestirse, cuidado de la apariencia personal, capacidad para caminar, capacidad para meterse y salir de la cama y capacidad para bañarse. Cada una de las funciones es valorada en una escala ordinal de tres grados que oscilan desde 0 (completamente incapaz) a 2 (sin la ayuda de nadie). La versión española ha seguido un riguroso proceso de adaptación transcultural con traducción, retrotraducción y estudio piloto³¹.

Adecuación conceptual y aplicabilidad

La tabla 1 resume los principales datos sobre la aplicabilidad y la adaptación transcultural de las diferentes versiones en castellano de las escalas analizadas^{12-14,26-29,31-65}.

El IB ha sido utilizado, en sus diferentes versiones, en cuatro de los seis contextos analizados. No se ha utilizado

Tabla 2. Resumen de los principales datos sobre fiabilidad de las diferentes versiones en castellano de las escalas de AVD

Instrumento	Fiabilidad		
	Consistencia interna	Prueba-reprueba	Equivalencia
Índice de Barthel, Mahoney y Barthel, 1965 ²⁰			
Baztán et al, 1993 ¹²	No informado	No informado	Kappa interobservador = 0,85. PC: 55/60 (autónomos frente a dependientes ⁴⁰) Acuerdo entre enfermería y cuidadores informales. Kappa: alimentación (0,31), lavarse (0,29), vestirse (0,39), arreglo (0,31), control deposiciones (0,17), IU (0,32), ir al retrete (0,46) y escalones (0,55). Para las AVD trasladarse y deambulación no obtuvieron kappa ³⁷
Índice de Katz, Katz et al, 1963 ²⁴			
Álvarez-Solar et al, 1992 ²⁶	α de Cronbach (enfermeras) = 0,86 y α de Cronbach (médicos) = 0,84 ²⁶	En una muestra de 38 personas, a las 2 semanas: Correlación intraobservador (Spearman) = 0,94 y Kendal-tau = 0,99 ²⁶	Acuerdo entre jueces, kappa = 0,84* (enfermeras frente a médicos) ²⁶
González-Montalvo et al, 1991 ²⁷	α de Cronbach = 0,95 y de 0,98 sin el ítem de incontinencia urinaria ⁶¹	No informado	Kappa interobservador entre 0,39 y 0,85 (el único superior a 0,70) ²⁷
Escala de incapacidad física de la Cruz Roja			
Guillén-Llera, 1986 ⁶²	No informado	No informado	Kappa interobservador entre 0,45 y 0,62 ²⁷

AVD: actividades de la vida diaria; PC: punto de corte.

* $p < 0,00001$.

ni en atención primaria ni en atención a domicilio. Principalmente se ha aplicado en instituciones de media estancia (unidades de convalecencia y rehabilitación). El IK, si bien ha sido utilizado en todos los contextos, aparece con mayor frecuencia en muestras de atención a domicilio. La EFCR no se ha utilizado en muestras comunitarias y se ha aplicado principalmente en atención especializada y atención a domicilio. La ACP-OARS sólo se ha aplicado en muestras comunitarias.

Todas las escalas se administraron mediante entrevista; el IB y el IK también se utilizaron como pruebas del desempeño (observación directa) y sólo el IB ha sido autoinformado. La fuente del informe fue en el caso del IB, en sus diferentes versiones, el paciente, el *proxy* (allegado) o ambos. Para el IK el paciente fue la principal fuente del informe, aunque en algunos casos también se preguntó a los allegados. En las escalas EFCR y ACP-OARS la fuente principal de información fueron tanto el paciente como el allegado.

El tiempo de administración y la cumplimentación de los ítems no se ha informado para ninguna de las escalas y sólo se ha informado de la existencia de entrenamiento y su duración para la versión del IK de Álvarez-Solar et al²⁶, donde los entrevistadores fueron entrenados por personal del equipo de atención primaria mediante sesiones de *role-playing*⁵⁶. En el resto de los casos, cuando se indica la ne-

cesidad de entrenamiento no se informa sobre éste. Los estudios tampoco han aportado datos sobre el desempeño (efectos suelo y techo) de ninguna de las escalas. En cuanto al significado de las puntuaciones, para las diferentes escalas, no existe consenso generalizado. En el caso del IB, los puntos de corte más frecuentes son 60^{50,54}, entre la dependencia moderada y leve, y 40, entre la dependencia moderada y severa^{36,52}. Para el IK, si bien existe una categorización preestablecida (7 categorías: A-G), en algunos casos ésta se agrupa en tres: independiente, parcialmente independiente y dependiente total, o bien como incapacidad leve, moderada y severa. Pero no todos los niveles equivalentes agrupan las mismas categorías^{27,59}. En otros estudios la interpretación de la escala se realiza en función del número de actividades en las que la persona es independiente⁶⁰. Los datos sobre la EFCR presentan cierto consenso, en tanto y en cuanto el punto de corte 3 parece ser el que mejor permite la discriminación entre la dependencia leve a moderada y la grave^{27,29,47,48}. Sobre el ACP-OARS no hay información al respecto. Para ninguna de ellas se han informado datos normativos.

El eje de aplicabilidad de las escalas ha sido sobre el que menor cantidad y calidad de información se ha aportado, lo que resulta significativo dado que esta información es esencial para su correcta administración clínica.

Propiedades psicométricas

La fiabilidad (tabla 2) se evaluó en tres de las cuatro escalas seleccionadas. Sobre el IB únicamente se informa de la fiabilidad por equivalencia para la versión de Baztán et al¹², cuya correlación interobservador supera los criterios de calidad mínimos a nivel grupal ($> 0,70$), pero no a nivel individual ($> 0,90$)⁶⁶. Dos de las versiones del IK presentan datos de fiabilidad, pero sólo la de Álvarez-Solar et al²⁶ presenta una información completa y obtiene coeficientes, en todos los casos, por encima de 0,70 y de test-retest por encima de 0,90. Sobre la versión de González-Montalvo et al²⁷ se informa un coeficiente α por encima de 0,90 y coeficientes kappa interobservador entre 0,39 y 0,85²⁷. En el caso de la EFCR, la fiabilidad sólo ha sido informada para una de las versiones⁶², y presenta valores de kappa inferiores a 0,70.

En cuanto a la validez, en todos los casos, excepto para la escala ACP-OARS, se han informado datos de validez de criterio predictivo (tablas 3 y 4). Sobre ninguna de las escalas se han informado datos de validez de criterio concurrente. En los análisis de utilidad predictiva, el IB resultó asociado a la mortalidad^{42,44,48}, a la institucionalización^{35,41,47,50,52,54}, a la utilización de servicios sociosanitarios^{35,39,47,48}, a resultados funcionales e indicadores de mejoría funcional^{32,33,36,38,43,44,50-52} y a las caídas⁴⁵. El IK se asoció a la mortalidad^{55,59}, la institucionalización⁶⁰ y a la utilización de servicios sociosanitarios⁵⁶. Y la EFCR resultó asociada a la mortalidad⁴⁸, la institucionalización^{47,48} y a la utilización de servicios sociosanitarios⁴⁷.

La validez de constructo se informó en todos los instrumentos. La estructura factorial sólo se ha analizado para la escala ACP-OARS³¹, que mostró un patrón unidimensional en un análisis factorial de todas las escalas del OARS. La validez discriminante, entre grupos conocidos, no se ha informado para ninguna de ellas.

Para las cuatro medidas se han llevado a cabo análisis de evidencias convergentes y divergentes (tabla 5). La obtención de evidencias convergentes se ha centrado en asociaciones con escalas de funcionamiento físico. El IB¹² se ha correlacionado con la escala Plutchik ($r = 0,85$)⁴⁰ y con la EFCR ($r = 0,86$)³⁵. La versión de González-Montalvo et al²⁷ y la de Cruz-Jentoft⁵⁷ del IK se asociaron a la EFCR con correlaciones entre 0,73 y 0,94²⁷ para la primera y de 0,87 para la última⁵⁸. La versión de la EFCR de Guillén-Llera⁶² se asoció, además, con el *Functional Autonomy Measurement System* (SMAF), con correlaciones entre 0,67 y 0,87⁶³. La obtención de evidencias divergentes se ha centrado en la búsqueda de asociaciones con escalas de otros dominios de CVRS, variables sociodemográficas y clínicas. Esta información se ha obtenido de estudios transversales o de análisis transversales en estudios longitudinales. La responsividad se evaluó formalmente para el IB en un estudio⁴¹ de la versión de Baztán et al¹². El trabajo, realizado en una unidad de convalecencia y rehabilitación, no espe-

cificó la intervención realizada y presentó un coeficiente de tamaño de efecto pequeño (0,31)⁶⁷.

Los instrumentos (tabla 6) que resultan con una valoración sumaria más alta son la versión del IB de Baztán et al¹² y la de Álvarez-Solar et al²⁶ del IK. Ambas son traducciones simples y sobre ellas se han presentado evidencias de su utilidad predictiva y convergencia o divergencia con otros instrumentos o constructos.

DISCUSIÓN

La revisión realizada nos ha permitido identificar cuatro escalas de actividades de la vida diaria: IB, IK, EFCR y ACP-OARS. Las evidencias obtenidas sobre éstas se han evaluado en función de tres ejes: adecuación conceptual, aplicabilidad y calidad psicométrica. En general, las evidencias son escasas y destaca la ausencia de datos sobre aplicabilidad e interpretabilidad. La información se ha centrado, mayoritariamente, en los aspectos psicométricos: de fiabilidad y validez.

La versión de Baztán et al¹² del IB y la versión del IK desarrollada por Álvarez-Solar et al²⁶ son los dos instrumentos sobre los que se ha obtenido mayor cantidad de información. Sin embargo, su proceso de validación lingüística ha sido deficiente; ambas son traducciones simples. La versión del IB se ha utilizado principalmente en atención especializada e instituciones de media y larga estancia como un instrumento de evaluación de intervenciones y ha presentado diferentes formas y fuentes de administración. El IK se ha utilizado en atención primaria y domiciliaria como un método de clasificación de pacientes o asignación de recursos (asistencia a domicilio) y se ha administrado únicamente mediante entrevista, y el paciente fue la fuente del informe en todos los casos. Para ninguno de ellos se han informado datos sobre el desempeño de los ítems (efecto suelo y techo), que sí han sido aportados en el ámbito internacional para el IK, donde se ha evidenciado un porcentaje elevado de efecto suelo en personas con baja discapacidad, lo que lo hace insensible a pequeños niveles de variación⁹. Tampoco se han obtenido datos normativos para ninguno de los dos. En relación con su aplicabilidad, sobre la versión de Baztán et al¹² no se ha obtenido información alguna y sobre la versión del IK²⁶ sólo ha sido mencionada en un estudio. Estas lagunas de información también se han evidenciado, en el ámbito internacional, en revisiones como la de McDowell y Newell⁹ o la de Sainsbury et al⁶⁸. En esta última se hace hincapié en la necesidad de dar respuesta a cuestiones tan relevantes como es la recomendación o no de entrenamiento o el uso de una guía de apoyo para la administración del IB en personas mayores.

Sobre la versión de Álvarez-Solar et al²⁶ del IK se ha presentado un número mayor de evidencias sobre fiabilidad, que superan en todos los casos los estándares de calidad para decisiones grupales; lo que no ocurre en el ámbito in-

Tabla 3. Resumen de los principales datos sobre utilidad predictiva de las diferentes versiones en castellano de las escalas de AVD: mortalidad, institucionalización y uso de recursos sociosanitarios

Instrumento	Utilidad predictiva		
	Mortalidad	Institucionalización	Utilización de recursos sociosanitarios
Índice de Barthel, Mahoney y Barthel, 1965 ²⁰			
Baztán et al, 1993 ¹²	Diferencia entre fallecidos y supervivientes en el IB al alta, los 6 meses ^c y al año ^{b44} Los fallecidos durante 6 meses presentaron un IB más bajo al alta (51,5/64,9 ^a) ⁴²	Los pacientes de residencias y unidades de larga estancia presentaron mayores grados de incapacidad funcional previa y al ingreso ³⁵ Diferencias en IB al ingreso entre los pacientes que regresaron a su domicilio o ingresaron en un centro residencial y los que fallecieron o reingresaron en el hospital ⁴¹	Estancia hospitalaria (16,7 en deteriorados frente a 10,6 en no deteriorados) ^a , máxima asociación en intervalos de 5 días o al establecer como PC: 15/16 con una OR = 5,5 (estancia \geq 15/16 días) ³³ Reutilización del servicio de urgencias al mes OR = 4,11 (deterioro funcional en AVD en consulta) ³⁹ MRM: días de estancia hospitalaria ^b en el IB actual y en el IB basal ³⁵ Mayor estancia hospitalaria OR = 4,29 ^a (IC del 95%, 2,81-6,56). PC < 45 (admisión) ⁴⁸ Necesidad de ayuda social a los 6 meses OR = 5,86 ^a (IB previo) OR = 3,18 ^a (ingreso). PC < 45 ⁴⁷
Alarcón-Alarcón et al, 1993 ⁴⁶	Hospitalaria OR = 7,25 (IC del 95%, 3,41-15,39); a 6 meses OR = 4,89 (IC del 95%, 3,29-7,26). PC < 65 (IB en admisión) ⁴⁸	Ingreso en residencia y ayuda domiciliaria OR = 10,7 ^a . PC < 45 (IB al ingreso) ⁴⁷	
Fontana-Fortuny et al, 2002 ⁴⁹	No informado	No informado	No informado
Índice de Barthel, Shah et al, 1989 ²²			
Cid y Damián, 1997 ¹³	No informado	IB < 60 o ingreso en residencia OR = 3,98 (IC del 95%, 1,62-9,79); muerte o institucionalización al año OR = 2,89 (IC del 95%, 0,98-8,52). PC < 90 (IB previo) ⁵⁰ Institucionalización OR = 30,7. PC < 40 (IB al alta) ⁵²	No informado
Índice de Barthel, Granger et al, 1979 ²³			
Grup de Treball, 1997 ⁵³	No informado	Retorno al domicilio OR ajustada = 3,67 ^b (PC \geq 60) y OR = 2,46 ^b (PC = 31-59). En la RL definitiva el IB resultó NS ⁵⁴	No informado
Índice de Katz, Katz et al, 1963 ²⁴			
Álvarez-Solar et al, 1992 ²⁶	RR = 1,23 (p = 0,511; categoría B) 5,32 ^b (categoría G) ⁵⁵	No informado	Probabilidad de no activo en el PADC (3 años), r/c la pérdida de 2 o más AVD (C o más) ^{a,56}
Cruz-Jentoft, 1991 ⁵⁷	No informado	No informado	No informado
Cruz-Jentoft, 1991 ¹⁴	Supervivencia: parcialmente dependientes OR = 1,18; totalmente dependientes OR = 3,67 ^b . En la RL la puntuación en IK fue NS ⁵⁹	El número de personas que fueron dadas de alta a domicilios se relacionó con un mejor IK en el ingreso. Destinos al alta de pacientes en función del IK al ingreso ^{c,60}	No informado
Fontana-Fortuny et al, 2002 ⁴⁹	No informado	No informado	No informado
González-Montalvo et al, 1991 ²⁷	No informado	No informado	No informado
Escala de incapacidad física de la Cruz Roja			
Guillén-Llera, 1972 ²⁸	Hospitalaria OR = 5,08. PC > 3 (antes de la enfermedad) ⁴⁸	Traslado a cuidado institucional OR = 3,41 ^b . PC > 3 (en admisión) ⁴⁸ Ingreso en residencia y previsión de ayuda social domiciliaria al alta OR = 3,41 ^b . PC > 3 ⁴⁷	La r con la estancia hospitalaria resultó NS ³⁵
Guillén-Llera, 1986 ⁶²	No informado	No informado	No informado
Cruz-Jentoft, 1991 ⁵⁷	No informado	No informado	No informado
Escala de actividades de cuidado personal del OARS, Fillenbaum, 1988 ³⁰			
Grau-Fibla et al, 1996 ³¹	No informado	No informado	No informado

AVD: actividades de la vida diaria; IB: índice de Barthel; IC: intervalo de confianza; MRM: modelo de regresión múltiple; NS: no significativo; OR: odds ratio; PADC: programa de atención a domicilio de crónicos; PC: punto de corte; r: correlación; r/c: relacionado con; RL: regresión logística; RR: riesgo relativo.

Deterioro funcional: una puntuación en el índice de Barthel inferior a la basal o previa a la consulta.

^ap \leq 0,05. ^bp \leq 0,01. ^cp \leq 0,001.

Tabla 4. Resumen de los principales datos sobre utilidad predictiva de las diferentes versiones en castellano de las escalas de AVD: resultados funcionales

Instrumento	Utilidad predictiva: resultados funcionales
Índice de Barthel, Mahoney y Barthel, 1965 ²⁰	
Baztán et al, 1993 ¹²	Deterioro funcional al mes de consulta OR = 12,8 (IC del 95%, 3,6-46,1). Pérdida funcional (IB previo-IB ingreso) ³² Deterioro funcional en AVD al mes del alta OR = 3,5 (IC del 95%, 2-6,3). IB al ingreso ³³ Riesgo de deterioro funcional al mes del alta ^b . Según el IBI, un nivel de incapacidad leve obtuvo una OR = 1, moderado OR = 6,43 y severo-total OR = 25. Estratificado: < 40 OR = 40,13; entre 55-40 OR = 10,32; entre 75-60 OR = 5,32 y entre 100-80 OR = 1 ³⁶ Capacidad de deambular al alta solo o con mínima ayuda OR = 1,02 (IC del 95%, 1,01-1,03). Puntuación previa en el IB (por punto) ³⁸ Indicadores mejoría funcional ^{*43} : 1. IB al alta r/c IB previo (0,41) ^d , con el IBI (0,71) ^d y con el deterioro funcional (-0,40) ^d 2. Ganancia funcional r/c IB previo (0,16) ^a y el deterioro funcional (0,25) ^a 3. IH r/c el IB previo (0,31) ^c y con IB al ingreso (-0,35) ^d 4. IHC r/c el IB al ingreso (0,31) ^c y el deterioro funcional (-0,37) ^d 5. E r/c el IB previo (0,15) ^a y con el deterioro funcional (0,18) ^a Ganancia funcional al año del alta. El IB previo ^b confirma su valor predictivo y el IB al ingreso adquiere un valor independiente ^{b,44} Incidencia de caídas en la hospitalización, el IB previo resultó predictor junto a otros predictores (OR = 1,013; IC del 95%, 1,002-1,024 ⁴⁵)
Alarcón-Alarcón et al, 1993 ⁴⁶	No informado
Fontana-Fortuny et al, 2002 ⁴⁹	No informado
Índice de Barthel, Shah et al, 1989 ²²	
Cid y Damián, 1997 ¹³	Al alta: no deambulación independiente OR = 4,5 (IC del 95%, 1,68-12,04), un I. Montebello < 0,5 OR = 0,26 (IC del 95%, 0,08-0,83) y un IB < 60 o ingreso en residencia 3,98 (IC del 95%, 1,62-9,79). A los 12 meses: no deambulación independiente OR = 9,90 (IC del 95%, 2,67-36,75). IB previo < 90 ⁵⁰ La RL ⁵¹ : 1. ganancia funcional ^b , estancia ^c y con eficiencia (NS). IB previo. 2. ganancia funcional ^c , estancia ^c y eficiencia (NS). IB al ingreso. El AM mostró que un IB previo < 60 (OR = 0,36 [IC del 95%, 0,14-0,93]) ^b y el IB al ingreso en cualquiera de sus valores: IB ≥ 60 (1), si 40-59 (OR = 3,71 [IC del 95%, 1,42-9,79]) ^b , si 20-39 (OR = 6,98 [IC del 95%, 2,49-19,56]) ^c y IB < 20 (OR = 7,14 [IC del 95%, 2,46-20,76]) ^c resultaron factores r/c la mejoría funcional. El IB al ingreso: si 40-59 (1), si 20-39 (OR = 0,28 [IC del 95%, 0,13-0,64]) ^b y < 20 (OR = 0,06 [IC del 95%, 0,02-0,13]) ^c , r/c IB al alta ≥ 60 ⁵²
Índice de Barthel, Granger et al, 1979 ²³	
Grup de Treball, 1997 ⁵³	No informado
Índice de Katz, Katz et al, 1963 ²⁴	
Álvarez-Solar et al, 1992 ²⁶	No informado
Cruz-Jentoft, 1991 ⁵⁷	No informado
Cruz-Jentoft, 1991 ¹⁴	No informado
Fontana-Fortuny et al, 2002 ⁴⁹	No informado
González-Montalvo et al, 1991 ²⁷	No informado
Escala de incapacidad física de la Cruz Roja	
Guillén-Llera, 1972 ²⁸	No informado
Guillén-Llera, 1986 ⁶²	No informado
Cruz-Jentoft, 1991 ⁵⁷	No informado
Escala de actividades de cuidado personal del OARS, Fillenbaum, 1988 ³⁰	
Grau-Fibla et al, 1996 ³¹	No informado

^ap ≤ 0,05. ^bp ≤ 0,01. ^cp ≤ 0,001. ^dp ≤ 0,0001. *Coeficiente de correlación de Pearson.

AM: análisis multivariado; AVD: actividades de la vida diaria; deterioro funcional: una puntuación en el índice de Barthel inferior a la basal o previa a la consulta; ganancia funcional: IB al alta - IB ingreso (IBA-IBI); I. Montebello: (IB alta - IB ingreso)/(IB previo - IB ingreso); índice de Heinemann (IH): IBA-IBI/10-IBIX100; índice de Heinemann corregido: 100 se sustituye por el valor real de IB previo; IC: intervalo de confianza; mejoría funcional: ganancia funcional mayor o igual a 20 puntos en el IB; NS: no significativo; OR: *odds ratio*; pérdida funcional: IB previo - IB ingreso (IBP-IBI); r/c: relacionado con; RL: regresión logística.

ternacional, donde ha sido catalogada como una escala con pocas evidencias de fiabilidad y validez^{9,69}. Para la versión de Baztán et al²⁶ del IB sólo se han presentado datos de fiabilidad por equivalencia. La escasa información que existe sobre la fiabilidad de esta escala, especialmente sobre su reproducibilidad test-retest, se ha puesto de manifiesto también en el ámbito internacional⁶⁸. Sin embargo,

en cuanto a la validez, esta versión del IB presenta mayor cantidad de evidencias sobre su utilidad predictiva y validez de constructo, y es la única sobre la que se han informado valores estandarizados de responsividad. Esto es paralelo, también, con lo que sucede en el ámbito internacional, donde las evidencias sobre la validez del IB son numerosas, tanto en relación con sus correlaciones

Tabla 5. Resumen de los principales datos sobre validez de constructo de las diferentes versiones en castellano de las escalas de AVD

Instrumento	Validez de constructo	
	Relaciones con otras escalas	Relaciones con otras variables
Índice de Barthel, Mahoney y Barthel, 1965 ²⁰		
Baztán et al, 1993 ¹²	r (EP) 0,85 ^a . PC (autonomía/dependencia): IB: 55/60 e EP: 4/5. Kappa intraobservador con EP = 0,75 y con IK = 0,77 ⁴⁰ r (IB ingreso-EFCR) 0,86 ^{b,35}	La relación con las variables sociodemográficas y el deterioro funcional resulta NS ³³ Se incluían en el deterioro funcional el grado de deterioro cognitivo ^c , el tipo de demencia ^a , la presencia de alucinaciones ^a , de apatía ^b y de depresión ^b , explicando el 25% de la variabilidad en la capacidad funcional ^{c,34} La r del IB ingreso con la edad 0,26 ^{b,35}
Alarcón-Alarcón et al, 1993 ⁴⁶	No informado	No informado
Fontana-Fortuny et al, 2002 ⁴⁹	No informado	Todos los ítems de IB al ingreso se correlacionaron con la edad (p = 0,000 y p = 0,016). El sexo resultó NS ⁴⁹
Índice de Barthel, Shah et al, 1989 ²²		
Cid y Damián, 1997 ¹³	No informado	El IB al ingreso y el IB previo se r/c la edad ^{b,52}
Índice de Barthel, Granger et al, 1979 ²³		
Grup de Treball, 1997 ⁵³	No informado	No informado
Índice de Katz, Katz et al, 1963 ²⁴		
Álvarez-Solar et al, 1992 ²⁶	No informado	Se encontró una diferencia significativa para la edad ^c y no con el sexo ²⁶ El deterioro del grado de autonomía (Katz) evoluciona con la pérdida de la capacidad cognitiva ^{a,56}
Cruz-Jentoft, 1991 ⁵⁷	r (EFCR) 0,87 ⁵⁸	No informado
Cruz-Jentoft, 1991 ¹⁴	No informado	No informado
Fontana-Fortuny et al, 2002 ⁴⁹	No informado	Todos los ítems se correlacionaron con la edad ^b y no con el sexo ⁴⁹
González-Montalvo et al, 1991 ²⁷	r Spearman (EFCR) 0,73 (médico 1); 0,90 (médico 2); 0,87 (enfermera) y 0,94 (auxiliar) ^{c,27}	No informado
Escala de incapacidad física de la Cruz Roja		
Guillén-Llera, 1972 ²⁸	r (IB) 0,86 ^b (p < 0,001) ³⁵ IB: S = 0,88; E = 0,78. Punto corte: 2 ²⁹	La r con la edad fue 0,21 ^{b,35}
Guillén-Llera, 1986 ⁶²	r (SMAF ingreso) 0,87; 0,83 (al mes); 0,67 (2 meses) y 0,76 (3 meses) ⁶³ r Spearman (IK) 0,73 (médico 1); 0,90 (médico 2); 0,87 (enfermera) y 0,94 (auxiliar de clínica) e IK: S entre 0,74 y 0,90; E entre 0,76 y 0,98. Tres niveles: leve (grados 0-I en la EFCR y A-B en el IK), moderado (grados II-III en la EFCR y C-D en el IK), grave (grados IV-V en la EFCR y E-G en el IK) ²⁷	No informado
Cruz-Jentoft, 1991 ⁵⁷	r Spearman (IK) 0,87 ⁵⁸	No informado
Escala de actividades de cuidado personal del OARS, Fillenbaum, 1988 ³⁰		
Grau-Fibla et al, 1996 ³¹	No informado	Mayor incapacidad para realizar la ACP se r/c deterioro cognitivo (no-leve frente a moderado-severo) OR = 6,49 (IC del 95%, 2,90-14,49) ^c , con la edad (60-64 frente a ≥ 75 años) OR = 3,23 (IC del 95, 2,16-4,83) ^c , con los síntomas psiquiátricos subjetivos (no leve frente a grave) OR = 2,49 (IC del 95%, 1,66-3,74) ^b , autopercepción del estado de salud (excelente/buena frente a regular/mala) OR = 2,41 (IC del 95%, 1,54-3,80) ^b , presencia de enfermedad crónica OR = 2,37 (IC del 95%, 1,01-5,54) ^a , número de visitas al médico/enfermera OR = 1,44 (IC del 95%, 1,04-2,00) ^a Cte = -4,2666 ^{c,65} La dependencia en ACP se asoció a un perfil en el EQ diferente del 1111 (mala calidad de vida) (OR = 2,5), con la existencia de problemas en las dimensiones actividades cotidianas (OR = 4,8) y movilidad (OR = 4,8) del EQ y en la de energía (OR = 1,9); dolor (OR = 2,9); movilidad física (OR = 3,4); reacciones emocionales (OR = 0,5) y aislamiento social (OR = 2,4) del PSN (todas excepto sueño) ⁶⁴

ACP: actividades de cuidado personal (AVD); AVD: actividades de la vida diaria; Cte: constante; E: especificidad; EFCR: escala de incapacidad física de la Cruz Roja; EP: escala Plutchik; EQ: EuroQol; IC: intervalo de confianza; NS: no significativa; OR: odds ratio; PSN: perfil de salud de Nottingham; r: correlación; r/c: relacionado con; S: sensibilidad.

^ap ≤ 0,05. ^bp < 0,001. ^cp < 0,0001.

Tabla 6. Valoración resumen sobre los principales indicadores de calidad de las diferentes versiones en castellano de las escalas de ABVD

Instrumento	Adaptación transcultural	Fiabilidad	Validez	Responsividad
Índice de Barthel, Mahoney y Barthel, 1965 ²⁰				
Baztán et al, 1993 ¹²	+	+	++	+
Alarcón-Alarcón et al, 1993 ⁴⁶	?	?	+	?
Fontana-Fortuny et al, 2002 ⁴⁹	?	?	+	?
Índice de Barthel, Shah et al, 1989 ²²				
Cid y Damián, 1997 ¹³	?	?	++	?
Índice de Barthel, Granger et al, 1979 ²³				
Grup de Treball, 1997 ⁵³	+	?	+	?
Índice de Katz, Katz et al, 1963 ²⁴				
Álvarez-Solar et al, 1992 ²⁶	+	++	++	?
Cruz-Jentoft, 1991 ⁵⁷	?	?	+	?
Cruz-Jentoft, 1991 ¹⁴	+	?	+	?
Fontana-Fortuny et al, 2002 ⁴⁹	?	?	+	?
González-Montalvo et al, 1991 ²⁷	+	+	+	?
Escala de incapacidad física de la Cruz Roja				
Guillén-Llera, 1972 ²⁸	NP	?	++	?
Guillén-Llera, 1986 ⁶²	NP	+	+	?
Cruz-Jentoft, 1991 ⁵⁷	NP	?	+	?
Escala de actividades de cuidado personal del OARS, Fillenbaum, 1988 ³⁰				
Grau-Fibla et al, 1996 ³¹	+++	?	+	?

ABVD: actividades básicas de la vida diaria; NP: no procede.

Criterios de valoración: 1. Proceso de adaptación transcultural (? : no informado; +: sólo traducción; ++: traducción-retrotraducción; +++: traducción-retrotraducción y prueba piloto); 2. Fiabilidad (? : no informada; +: alguna aspecto de fiabilidad evaluado (1 o 2); ++: todos los aspectos examinados en un único estudio; +++: todos los aspectos examinados en varios estudios); 3. Validez (? : no informada; +: un único aspecto de validez evaluado; ++: varios aspectos de validez evaluados; +++: todos los aspectos examinados); 4. Responsividad (? : no informada; +: evidencia procedentes de un único estudio; ++: evidencias procedentes de más de un estudio).

con otras escalas como sobre su capacidad predictiva de resultados adversos^{9,69}.

En cuanto al resto de escalas, la escala ACP-OARS presenta un adecuado proceso de adaptación transcultural, pero llevado a cabo como subescala del cuestionario OARS, y la EFCR se ha creado en castellano, lo que limita su comparabilidad con el ámbito internacional. Las evidencias obtenidas sobre su aplicabilidad y sus propiedades psicométricas son, para ambas, pocas o muy débiles. Cabría destacar que la primera se ha empleado en muestras comunitarias, con la finalidad de identificar necesidades de servicios o asignar recursos y la EFCR, básicamente, se ha utilizado para valorar cronicidad e invalidez desde un punto de vista médico.

Globalmente, las escalas de AVD identificadas en la literatura médica española adolecen de problemas en su adecuación conceptual, apareciendo un elevado número de versiones de una misma escala con procesos de adaptación transcultural poco rigurosos y sin información sobre el desempeño de sus ítems. Son escalas que, en su mayoría, se han aplicado en más de un contexto, excepto la ACP-OARS, pero con escasas pruebas de su calidad en todos

ellos (en nuestro país). El IB es el más utilizado en instituciones de media (unidades de convalecencia y rehabilitación) y larga estancia; el resto se utiliza más para la planificación de necesidades de cuidados de larga duración desde contextos clínicos o comunitarios.

La aplicabilidad es el apartado más deficiente: no se han informado datos normativos, puntos de corte estandarizados ni normas o instrucciones de administración. Los criterios psicométricos son el eje con mayor cantidad de información reunida, pero hay notables lagunas. Los datos sobre fiabilidad test-retest y consistencia interna son escasos; la equivalencia interobservador es el aspecto más estudiado de la fiabilidad. No se han obtenido evidencias de validez de criterio concurrente ni análisis confirmatorios sobre la estructura de los instrumentos, pero sí un elevado número de evidencias sobre su utilidad predictiva y convergencia/divergencia. Es destacable, asimismo, la ausencia de datos sobre responsividad para la mayor parte de las escalas, salvo el IB.

Esta distribución de hallazgos psicométricos no es casual. En consonancia con el hecho de que las medidas de AVD aportan una información muy valiosa como índices pro-

nósticos de resultados adversos³, la mayoría de los estudios de validez encontrados han sido de utilidad predictiva. Estas escalas tienen un fuerte efecto techo (o suelo, si el sentido de la puntuación es de mayor a menor funcionalidad) en la investigación clínica y las convierte en poco útiles para la monitorización clínica de pacientes y la evaluación de resultados⁶. Consecuentemente, no se han encontrado apenas estudios de responsividad, salvo en el caso del IB. A primera vista, resulta llamativo los pocos estudios analítico-factoriales y de consistencia interna realizados con estas escalas, pero esta metodología asume la existencia de un dominio infinito de ítems que sean indicadores de un concepto abstracto o rasgo latente⁷⁰. Esta lógica no está detrás de las escalas de AVD: los ítems son indicadores de un conjunto muy limitado de actividades o roles, y el universo de medida se agota en ellos; además, en su génesis ha primado más lo pragmático que lo conceptual^{6,9}.

Para finalizar el examen crítico, dos cuestiones generales. Por un lado, la ausencia de estudios que comparen sistemáticamente la aplicabilidad y calidad psicométrica de dos o más escalas simultáneamente. Por otro, los instrumentos revisados son viejos, contruidos antes de la Clasificación Internacional de Deficiencias, Discapacidades y Minusvalías⁷¹, y consiguientemente antes de la Clasificación Internacional del Funcionamiento, de la Discapacidad y de la Salud⁷² y de otros modelos conceptuales relevantes, como el proceso de discapacidad⁷³. En el ámbito internacional, existe ya una nueva generación de instrumentos de estatus funcional con una sólida validez de contenido y propiedades psicométricas muy robustas, examinadas con metodología de "teoría de la respuesta al ítem" (TRI), como el Activity Measure for post-Acute Care (AM-PAC)⁷⁴.

Si bien los datos son contundentes, es necesario tener en cuenta las limitaciones derivadas de los criterios de selección establecidos, ya que esta revisión no ha incluido las medidas genéricas de salud ni otras medidas multidimensionales. Sin embargo, estos perfiles no disponen de escalas específicas de AVD, aunque algunos de ellos tienen medidas monoítem de AVD, como las láminas COOP/WONCA⁷⁵ o el EuroQol 5D⁷⁶. Además, su uso específico con personas mayores es raro. Una posible excepción es la escala de funcionamiento físico del SF-36⁷⁷, pero sólo uno de sus ítems se refiere a AVD. Otra segunda limitación es inherente al empleo de búsquedas electrónicas y recuperación de documentos. En nuestra estrategia ha primado la sensibilidad; para ello, se han empleado múltiples sinónimos de los términos de búsqueda y múltiples bases de datos, y las búsquedas electrónicas se han completado con búsquedas manuales y rastreo de referencias. Una tercera limitación deriva del hecho de no haber considerado para su inclusión la literatura médica gris.

Los hallazgos de esta revisión resultan paradójicos si tenemos en cuenta la relevancia y frecuente utilización de estas medidas, especialmente el IB y el IK, tanto en la investigación como en la clínica. La única explicación posi-

ble radica en la aceptación general de ciertas escalas por su reconocimiento y uso generalizado más que por la evidencia demostrada de su calidad psicométrica⁹. Finalmente, recomendamos: a) la realización de trabajos dirigidos a la estandarización de la aplicación de estos instrumentos y b) la valoración concurrente de varios de ellos, pero también c) el desarrollo o adaptación de nuevos instrumentos, mejor fundados conceptualmente y metodológicamente basados en modelos psicométricos más potentes (como los modelos de Rasch), que ya han demostrado su utilidad y viabilidad en este campo^{9,78,79}.

BIBLIOGRAFÍA

1. Spector WD, Fleishman, JA. Combining activities of daily living with instrumental activities of daily living to measure functional disability. *J Gerontol B Psychol Sci Soc Sci.* 1998;53:S46-S57.
2. Rodgers W, Miller B. A comparative analysis of ADL questions in surveys of older people. *J Gerontol B Psychol Sci Soc Sci.* 1997;52:21-36.
3. Guralnik JM, Fried LP, Salive ME. Disability as a public health outcome in the aging population. *Annu Rev Public Health.* 1996;17:25-46.
4. Avlund K. Methodological challenges in measurements of functional ability in gerontological research. A review. *Aging (Milano).* 1997;9:164-74.
5. Wiener JM, Hanley RJ, Clark R, Van Nostrand JF. Measuring the activities of daily living: comparisons across national surveys. *J Gerontol.* 1990;45:S229-37.
6. Coster WJ, Haley SM, Andres PL, Ludlow LH, Bond TLY, Ni P. Refining the conceptual basis for rehabilitation outcome measurement. *Med Care.* 2004;42 Suppl 1:162-72.
7. Lawton MP, Brody EM. Assessment of older people: self-maintaining and instrumental activities of daily living. *Gerontologist.* 1969;9:179-86.
8. Katz S. Assessing self-maintenance: Activities of daily living, mobility, and instrumental activities of daily living. *J Am Geriatr Soc.* 1983;31:721-27.
9. McDowell I, Newell C. *Measuring health: a guide to rating scales and questionnaires.* 2.ª ed. New York: Oxford University Press; 1996.
10. Haywood KL, Garratt AM, Schmidt L, Mackintosh AE, Fitzpatrick R. Health status and Quality of Life in older people: a review. Report from the Patient-assessed Health Instruments Group to the Department of Health, April 2004. Disponible en: <http://phi.uhce.ox.ac.uk>
11. Badia X, Salamero M, Alonso J. La medida de la salud. Guía de escalas de medición en español. 3.ª ed. Barcelona: Edimac; 2002.
12. Baztán JJ, Pérez-del-Molino J, Alarcón T, San Cristóbal E, Izquierdo G, Manzarbeitia J. Índice de Barthel: instrumento valido para la valoración funcional de pacientes con enfermedad cerebrovascular. *Rev Esp Geriatr Gerontol.* 1993;28:32-40.
13. Cid-Ruzafa J, Damián-Moreno J. Valoración de la discapacidad física: El Índice de Barthel. *Rev Esp Salud Publica.* 1997;71:127-37.
14. Cruz-Jentoft AJ. El índice de Katz. *Rev Esp Geriatr Gerontol.* 1991;26:338-48.
15. McDonald SJ, Lefebvre C, Clarke MJ. Identifying reports of controlled trials in the BMJ and the Lancet. *BMJ.* 1996;313:1116-7.
16. Scientific Advisory Committee of the Medical Outcomes Trust. Assessing health status and quality-of-life instruments: attributes and review criteria. *Qual Life Res* 2002;11:193-205.
17. Valderas J, Ferrer M, Alonso J. Instrumentos de medida de calidad de vida relacionada con la salud y de otros resultados percibidos por los pacientes. *Med Clin (Barc).* 2007;125 Supl 1:56-60.
18. Muñiz-Fernández J. *Teoría Clásica de los test.* Madrid: Pirámide; 1992.
19. Acquadro C, Conway K, Giroulet C, Mear I. *Linguistic validation manual for Patient-Reported Outcomes (PRO) Instruments.* Lyon: MAPI Research Institute; 2004.
20. Mahoney FI, Barthel DW. Functional evaluation: the Barthel Index. *Med State Med J.* 1965;14:61-5.
21. Collin C, Wade DT, Davies S, Horn V. The Barthel ADL Index: a reliability study. *Int Disabil Stud.* 1988;10:61-3.

22. Shah S, Vanclay F, Cooper B. Improving the sensitivity of the Barthel Index for stroke rehabilitation. *J Clin Epidemiol.* 1989;42:703-9.
23. Granger CV, Albrecht GL, Hamilton BB. Outcome of comprehensive medical rehabilitation: measurement by PULSES profile and the Barthel Index. *Arch Phys Med Rehabil.* 1979;60:145-54.
24. Katz S, Ford A, Moskowitz RW, Jackson BA, Jaffe MW. Studies of illness in the aged. The Index of ADL: a standardized measure of biological and psychosocial function. *JAMA.* 1963;185:914-9.
25. Casiano E, Paddon D, Ostir G, Sheffield M. Assessing functional status measures in older adults: a guide for healthcare professionals. *Phys Ther Rev.* 2002;7:89-101.
26. Álvarez-Solar M, Alaiz-Rojo AT, Bru-Gurpegui E, Cabañeros-Vicente JJ, Calzón-Frecho M, Cosío-Rodríguez I, et al. Capacidad funcional de pacientes mayores de 65 años, según el índice de Katz. Fiabilidad del método. *Aten Primaria.* 1992;10:812-6.
27. González-Montalvo JI, Rodríguez-Pascual C, Diestro-Martín P, Casado-Casado MT, Vallejo-Pezuela MI, Calvo-Catalán MJ. Valoración funcional: comparación de la escala de Cruz Roja con el índice de Katz. *Rev Esp Geriatr Gerontol.* 1991;26:197-202.
28. Guillén-Llera F, García-Antón A. Ayuda a domicilio. Aspectos médicos en Geriatría. *Rev Esp Geriatr Gerontol.* 1972;7:339-46.
29. Regalado-Doña P, Valero-Ubierna C, González-Montalvo JI, Salgado-Alba A. La escalas de la Cruz Roja veinticinco años después: estudio de la validez en un servicio de Geriatría. *Rev Esp Geriatr Gerontol.* 1997;32:93-9.
30. Fillenbaum GG. *Multidimensional Functional Assessment of Older Adults.* The Duke Older Americans Resources and Services Procedures. Hillsdale: Lawrence Erlbaum Associates Publishers; 1988.
31. Grau-Fibla G, Eiroa-Patiño P, Cayuela-Domínguez A. Versión española del OARS Multidimensional Functional Assessment Questionnaire: adaptación transcultural y medida de la validez. *Aten Primaria.* 1996;17:486-95.
32. Gutiérrez-Rodríguez J, Alonso-Álvarez M, Valera-Suárez C, Solano-Jaurrieta JJ. Factores de riesgo asociados a deterioro funcional en el anciano tras su consulta en un servicio de urgencias. *Rev Mult Gerontol.* 1999;9:139-46.
33. Gutiérrez-Rodríguez J, Alonso-Álvarez M, Varela-Suárez C, Solano-Jaurrieta JJ. Deterioro funcional en el anciano hospitalizado por insuficiencia cardíaca. Análisis de incidencia y los factores de riesgo. *Rev Mult Gerontol.* 1998;8:221-8.
34. Artaso-Irigoyen B, Goñi-Sarriés A, Gómez-Martínez AR. Deterioro funcional del paciente con demencia: influencia del grado de deterioro cognitivo, la comorbilidad física y las alteraciones neuropsiquiátricas. *Rev Esp Geriatr Gerontol.* 2002;37:210-5.
35. González-Montalvo JI, Alarcón T, Del Río M, Gotor P, Aceituno V. Resultados asistenciales de un equipo consultor geriátrico en un hospital general. Cuatro años de actividad. *Rev Esp Geriatr Gerontol.* 1999;34:336-45.
36. Gutiérrez-Rodríguez J, Galeano-Arbolea R, Alonso-Álvarez M, Varela-Suárez C, Solano-Jaurrieta JJ. ¿Es útil el índice de Barthel en la detección de riesgo de deterioro funcional en pacientes ancianos ingresados por enfermedad aguda? *Rev Mult Gerontol.* 1998;8:213-9.
37. Artaso-Irigoyen B, Goñi-Sarriés A, Gómez-Martínez AR, García-Nicolás M. Valoración funcional directa e indirecta del paciente con demencia. *Geriatría.* 2002;18:35-9.
38. González-Montalvo JI, Alarcón T, Sáez P, Barcena A, Gotor P, Del Río M. La intervención geriátrica puede mejorar el curso clínico de los ancianos frágiles con fractura de cadera. *Med Clin (Barc).* 2001;116:1-5.
39. Gutiérrez-Rodríguez J, Varela-Suárez C, Alonso-Álvarez M, Solano-Jaurrieta JJ. Deterioro funcional en actividades básicas de la vida diaria tras la consulta en un servicio de urgencias. *An Med Interna.* 2000;17:238-42.
40. Perlado F, Alastuey C, Espinosa C, Clerencia M, Ochoa P, Domingo D. Valoración del estado funcional en ancianos hospitalizados: la escala Plutchik. *Med Clin (Barc).* 1996;107:45-9.
41. Diestre G, Villarmou MT. Evaluación de la capacidad funcional y características de los pacientes ingresados en la Unidad de Convalecencia y Rehabilitación (UCR). *Rev Mult Gerontol.* 2000;10:26-32.
42. Villarmou MT, Diestre G, Guirao M. Evaluación de los pacientes dados de alta con el diagnóstico de accidente vasculocerebral de una Unidad de Convalecencia y Rehabilitación (UCR). *Rev Mult Gerontol.* 2000;10:33-7.
43. Valderrama-Gama E, Baztán-Cortés JJ, Molpeceres-García-del-Pozo J, Pérez-Muñoz J, Fernández-Adarve MM, et al. Evaluación de la mejoría funcional y la estancia en una unidad de rehabilitación geriátrica. *Rev Esp Geriatr Gerontol.* 1997;32:200-6.
44. Valderrama-Gama E, Molpeceres-García-del-Pozo J. Seguimiento a largo plazo de los pacientes ingresados en una unidad de media estancia o convalecencia. *Rev Esp Geriatr Gerontol.* 1998;33:67-72.
45. Miñana-Climent JC, San Cristóbal-Velasco E, Arche-Coto JM, Rodríguez-Piñera MA, Fernández-Fernández M, Fernández-Fernández M. Características y factores asociados a caídas en pacientes con enfermedad cerebrovascular. *Rev Esp Geriatr Gerontol.* 2005;40:24-30.
46. Alarcón-Alarcón MT, González-Montalvo JI, Salgado-Alba A. Valoración funcional del paciente anciano. En: Salgado A, Alarcón MT, editores. *Valoración del paciente anciano.* Barcelona: Masson; 1993. p. 47-72.
47. Alarcón-Alarcón T, González-Montalvo JI, Bárcena-Álvarez A, Salgado-Alba A. Estudio predictivo de las necesidades sociales al alta hospitalaria y a los seis meses en ancianos ingresados por enfermedad aguda. *Rev Esp Geriatr Gerontol.* 1998;33:209-18.
48. Alarcón T, Bárcena A, González-Montalvo JI, Penalosa C, Salgado A. Factors predictive of outcome on admission to an acute geriatric ward. *Age Ageing.* 1999;28:429-32.
49. Fontana-Fortuny C, Estany-Almirall J, Pujol-Salud J, Segarra-Solares I, Jordán-Gómez Y. Concordancia entre índices de dependencia en las actividades de la vida diaria. Experiencia de aplicación en la población geriátrica de ámbito rural. *Enferm Clin.* 2002;12:47-53.
50. Baztán JJ, Fernández-Alonso M, Aguado R, Socorro A. Resultados al año de la rehabilitación tras fractura de fémur proximal en mayores de 84 años. *An Med Interna.* 2004;21:433-40.
51. Baztán JJ, Domenech JR, González M, Forcano S, Morales C, Ruipérez I. Ganancia funcional y estancia hospitalaria en la unidad de geriatría de media estancia del Hospital Central de la Cruz Roja de Madrid. *Rev Esp Salud Pública.* 2004;78:355-66.
52. Baztán JJ, González M, Morales C, Vázquez E, Morón N, Forcano S, et al. Variables asociadas a la recuperación funcional y la institucionalización al alta en ancianos ingresados en una unidad geriátrica de media estancia. *Rev Clín Esp.* 2004;204:574-82.
53. Grup de Treball sobre Rehabilitació Geriàtrica de la Societat Catalana Balear de Geriatria i Gerontologia. *Aspectes generals de la rehabilitació geriàtrica.* Barcelona: Glosa Edicions; 1997.
54. Sarbatés O, Miralles R, Ferre M, Esperanza A, García-Palleiro P, Llorach I, et al. Factores predictivos de retorno al domicilio en pacientes ancianos hospitalizados. *An Med Interna.* 1999;16:407-14.
55. Segura-Noguera JM, Bastida-Bastús N, Martí-Guadano N, Riba-Romera M. Enfermos crónicos domiciliarios. Análisis descriptivo y de supervivencia a los dos años de la implantación de un programa de atención domiciliaria. *Aten Primaria.* 1997;19:351-6.
56. Segura-Noguera JM, Barreiro-Montaña MC, Bastida-Bastús N, Collado-Gastalver D, González-Solanellas M, Linares-Juan-de-Sentmenat C, et al. Enfermos crónicos domiciliarios: valoración integral psicocognitiva y caídas durante 3 años de seguimiento. *Aten Primaria.* 2000;25:35-43.
57. Cruz-Jentoft AF. Evaluación de la función física en el anciano. *Medicine.* 1991;86:3326-31.
58. Sanz-Julían C, Morente-Páez M, Calvo-Villaverde C, Gallego-Casado P, Carra-Meana B. Valoración psicofísica del anciano. Escalas utilizadas. *Rev Rol Enferm.* 1995;XVIII:17-20.
59. Palomoa L, Gervas J. Mortalidad a los dos años en pacientes crónicos confinados en el domicilio. *Aten Primaria.* 2000;25:106-14.
60. Salvá A, Aloy J, Castro D, Faustino A, Fabregas N, Petit TM, et al. La evaluación geriátrica en una unidad de convalecencia. Análisis descriptivo de la actividad de un año. *Rev Esp Geriatr Gerontol.* 1996;31:205-10.
61. Iturría JA, Márquez S. Situación de los ancianos tras un episodio de hospitalización: estado de salud, recursos materiales y apoyo social. *Rev Gerontol.* 1996;6:92-7.
62. Guillén-Llera F. Unidades geriátricas de hospitalización. En: Salgado Alba A, Guillén Llera F, Díaz de la Peña J, editors. *Tratado de geriatría y asistencia geriátrica.* Barcelona: Salvat S.A.; 1986. p. 59-67.
63. Gamboa B, Morlanes T, Galindo J, Cucullo JM, Calvo MJ. Uso de las escalas de actividades de la vida diaria en un hospital de día geriátrico. *Rev Gerontol.* 1992;2:117-20.

64. Aspiazu-Garrido M, Cruz-Jentoft A, Villagrasa-Ferrer Jr, Abanades-Herranz JC, García-Marín N, Alvear-Valero-de-Bernabé FA. Factores asociados a mal estado de salud percibido o a mala calidad de vida en personas mayores de 65 años. *Rev Esp Salud Publica*. 2002;76:683-99.
65. Espejo-Espejo J, Martínez-de-la-Iglesia J, Aranda-Lara JM, Rubio-Cuadarado V, Enciso-Bergé I, Zunzunegui-Pastor MV, et al. Capacidad funcional en mayores de 60 años y factores sociosanitarios asociados. *Aten Primaria*. 1997;20:3-11.
66. Cronbach L. Fundamentos de los test psicológicos. Madrid: Biblioteca Nueva S.L.; 1998.
67. Cohen J. Statistical power analysis for the behavioral sciences. New York: Academic Press; 1977.
68. Sainsbury A, Seebass G, Bansal A, Young JB. Reliability of the Barthel Index when used with older people. *Age Ageing*. 2005;34:228-32.
69. Bowling A. La medida de la salud. Revisión de las escalas de medida de la calidad de vida. Barcelona: Masson S.A.; 1994.
70. Nunnally JC, Bernstein IB. Teoría psicométrica. México: McGraw-Hill; 1995.
71. World Health Organization. International Classification of Impairments, Functioning, Disabilities and Handicaps. Geneva: WHO; 1980.
72. World Health Organization. International Classification of Functioning, Disability and Health: ICF. Geneva: WHO; 2001.
73. Verbrugge LM, Jette AM. The disablement process. *Soc Sci Med*. 1994;38:1-14.
74. Haley SM, Coster WJ, Andres PL, Ludlow LH, Ni P, Bond TLY, et al. Activity outcome measurement for postacute care. *Med Care*. 2004;42 Suppl 1:149-161
75. Scholten JHG, Van Weel C. Functional status assessment in family practice. The Dartmouth COOP Functional Health Assessment Charts/WONCA. Lelystad: Meditekst; 1992. p. 17-51.
76. Euroqol Group. Euroqol-a new facility for the measurement of health related quality of life. *Health Policy*. 1990;16:199-208.
77. López-García E, Banegas JR, Pérez-Regadera AG, Gutiérrez-Fisac JL, Alonso J, Rodríguez-Artalejo F. Valores de referencia de la versión española de Cuestionario de Salud SF-36 en población adulta de más de 60 años. *Med Clin (Barc)*. 2003;120:568-73.
78. Avlund K, Kreiner S, Schultz-Larsen K. Construct validity and the Rasch model: functional ability of health elderly people. *Scand J Soc Med Suppl*. 1993;21:233-46.
79. McHorney C, Cohen A. Equating health status measures with the Item Response Theory: Illustrations with functional status items. *Med Care*. 2000;38 Suppl 9:1143-59.