

# Concordancia interobservador en la valoración funcional mediante el índice de Barthel

F. Formiga, C. Ortega, C. Cabot, F. Porras, J. Mascaro y R. Pujol

UFISS Geriatría-Servicio de Medicina Interna. Hospital Universitari de Bellvitge. L'Hospitalet de Llobregat. Barcelona.

**Fundamentos.** El objetivo del estudio es evaluar la reproductividad interobservador entre enfermeras clínicas expertas en geriatría y médicos residentes en su primer contacto con el índice de Barthel (IB).

**Material y métodos.** Se evaluaron 100 pacientes. El IB, dos semanas previas al ingreso, se recogió mediante entrevista al enfermo o al cuidador principal. Se excluyeron 18 pacientes en que fue imposible la revaloración.

**Resultados.** En los 82 pacientes incluidos la media del IB valorada por la enfermera fue de 87,3 (19) y la del médico residente de 88,2 (21). En 40 pacientes existía alguna diferencia en el valor medio del IB. El acuerdo al examinar el total de la escala era bueno ( $r = 0,793$ ); no obstante, el acuerdo interobservador era bajo ( $\kappa < 0,4$ ) en los apartados de comer, arreglo y trasferencias.

**Conclusiones.** La baja reproductividad interobservador en algunos apartados del IB entre enfermeras clínicas expertas en geriatría y médicos residentes en su período de formación en geriatría sugiere la necesidad que tienen éstos de un período de aprendizaje en evaluación de la capacidad.

**PALABRAS CLAVE:** Índice de Barthel, anciano, reproductividad.

Formiga F, Ortega C, Cabot C, Porras F, Mascaro J, Pujol R. Concordancia interobservador en la valoración funcional mediante el índice de Barthel. Rev Clin Esp. 2006;206(5):230-2.

## Introducción

La valoración funcional se utiliza para identificar el grado de independencia que alcanza un individuo en las actividades de la vida diaria (AVD) y en la deambulación. En función de su complejidad se distingue entre actividades básicas e instrumentales de la vida diaria<sup>1</sup>.

Las AVD básicas (ABVD) incluyen actividades del autocuidado personal como el baño, aseo, comer, vestirse, continencia y movilidad, entre otros. Existen varias escalas validadas para valorar las ABVD, una de las más utilizadas internacionalmente es el índice de Barthel (IB)<sup>1-5</sup>. El objetivo del presente trabajo era estudiar la reproductividad interobservador entre enfermeras clínicas

Interobserver concordance in functional assessment by Barthel index

**Background.** The aim of this study is to analyze interobserver «reproductiveness» between clinical nurses in gerontology and residents in family medicine in their first contact with the Barthel Index (BI).

**Material and methods.** One hundred patients were assessed. BI at two weeks before admission was scored by a patient or carer interview. Eighteen patients were excluded because interviews were not completed.

**Results.** Mean BI of 82 patient included, assessed by nurses was 87.3 and by residents 88.2. In 40 patients some differences in the mean values of BI were observed. When analyzing the overall BI score the agreement was high ( $r = 0.793$ ), but interobserver agreement was low ( $\kappa < 0.4$ ) in some fields such as eating, dressing and transfers and medium in others ( $\kappa$  from 0.40 to 0.75).

**Conclusions.** Low reproductiveness interobserver in some patients of the BI when clinical nurses in gerontology and residents in family medicine are compared suggest that they need specific training in functional capacity evaluation in order to improve the BI use.

**KEY WORDS:** Barthel index, elderly, reproductiveness.

expertas en geriatría y médicos de familia en formación en geriatría en su primer contacto con el IB. La hipótesis era que aunque el IB sea un instrumento sencillo de utilizar la reproductividad que se observaría no sería excelente debido a que probablemente para la óptima utilización del IB es importante un período de aprendizaje.

## Pacientes y métodos

Se estudiaron 100 pacientes ingresados durante el año 2003 que procedían de la comunidad. El IB<sup>2</sup>, de dos semanas previas al ingreso, se recogió mediante entrevista al enfermo o al cuidador principal por duplicado con una diferencia de 48 horas en las evaluaciones (para que no se recuerden las respuestas se acepta como válido dicho intervalo). La entrevista fue realizada, por una parte, por dos enfermeras clínicas expertas en geriatría y por otra parte por un equipo de médicos residentes, en el que era su primer contacto con el IB, y a los que se les explicó brevemente cómo utilizarlo. En 50 de los pacientes inició la entrevista la enfermera clínica y en los otros 50 pacientes el médico re-

Correspondencia: F. Formiga.  
UFISS Geriatría. Servicio de Medicina Interna.  
Hospital Universitari de Bellvitge.  
C/ Feixa Llarga s/n.

08907 L'Hospitalet de Llobregat. Barcelona.  
Correo electrónico: fformiga@csuib.scs.es

Aceptado para su publicación el 5 de noviembre de 2004.

sidente. Se excluyeron 18 pacientes: 11 pacientes en los que ocurrió un nuevo evento que imposibilitó repetir la historia al paciente o que fue imposible localizar al mismo cuidador y 7 pacientes debido a que fallecieron o fueron dados de alta antes de realizarse la segunda evaluación. En un muestra previa con 30 pacientes se había demostrado una buena acuerdo interobservador entre las dos enfermeras clínicas que realizaron las entrevistas, ya sea en la fiabilidad total (correlación de Spearman;  $r = 0,98$ ) o en los diversos campos del IB (índice de kappa superior al 0,89 en todos ellos). Se evaluó el nivel cognitivo de los pacientes con el *Short Portable Mental Status Questionnaire* (SPMSQ)<sup>6</sup>.

En los pacientes con más de 4 fallos se realizó la entrevista al cuidador principal (siempre al mismo), así como también en los pacientes con problemas de comunicación (afasia, disartria) o afectos de un cuadro confusional agudo.

El IB es una escala ordinal que incluye 10 áreas de actividades básicas de la vida diaria (comer, bañarse, vestirse, aseo personal, continencia urinaria, continencia fecal, uso del retrete, traslado sillón/cama, deambulación y uso de escaleras). La puntuación total varía de 0-100 (anexo 1). Se ha confirmado la validez, sensibilidad y reproductividad intra-observador<sup>7-9</sup> del IB y también una buena reproductividad interobservador<sup>9-10</sup>. Además el IB puede ser utilizado por entrevistadores no sanitarios entrenados<sup>11</sup>, bastando al parecer una sesión única de explicación para aprender a puntuar el IB<sup>12</sup>. Para valorar la complejidad de los pacientes se aplicó el índice de comorbilidad de Charlson<sup>13</sup>.

Los resultados se expresan con la media con desviación estándar (DE); la significación estadística de las diferencias se valoró mediante el test de la «t» de Student para variables cuantitativas y test no paramétrico en caso de distribución no normal de la variable. La prueba de la Chi cuadrado con la

prueba exacta de Fisher se utilizaron en la comparación de variables cualitativas. Para evaluar el acuerdo interobservador en los diversos apartado del IB se utilizó la prueba kappa de Cohen. Un valor kappa > 0,75 se consideró indicativo de acuerdo excelente, entre 0,40-0,75 acuerdo aceptable y < 0,4 acuerdo pobre<sup>14</sup>. La fiabilidad de la puntuación total de la escala se evaluó mediante el coeficiente de correlación de Spearman. Los resultados se consideraron significativos cuando  $p < 0,05$ .

## Resultados

De los 82 pacientes evaluados en el estudio, 45 eran mujeres (55%) y 37 varones con una edad media de 72,6 (11) años (rango: 40-96 años). El valor medio del índice de Charlson fue de 2,6 (2,2) (rango: 0-10). En 21 pacientes (25%) se realizó la entrevista al cuidador. Todos los pacientes provenían de su domicilio; 24 pacientes vivían solos (29%). Treinta y siete pacientes estaban casados (45%), 33 pacientes (40%) eran viudos, 8 pacientes (10%) eran solteros y 4 pacientes (5%) estaban separados. Veintiún pacientes (26%) no tenían estudios, 47 pacientes (57%) tenían estudios primarios, 11 (13%) pacientes estudios secundarios y 3 pacientes (4%) eran universitarios.

### Evaluación del IB previo al ingreso

Los valores medios del IB previo al ingreso valorados por la enfermera eran de 87,3 (19), existiendo una buena concordancia ( $r = 0,793$ ) con los valores me-

## ANEXO 1 Índice de Barthel

Comida	10	Independiente. Capaz de comer por sí solo en un tiempo razonable. La comida puede ser cocinada servida por otra persona
	5	Necesita ayuda para cortar la carne, extender la mantequilla..., pero es capaz de comer solo
Lavado (baño)	0	Dependiente. Necesita ser alimentado por otra persona
	5	Independiente. Capaz de lavarse entero, de entrar y salir del baño sin ayuda y de hacerlo sin que una persona supervise
Vestido	0	Dependiente. Necesita algún tipo de ayuda o supervisión
	10	Independiente. Capaz de ponerse y quitarse la ropa sin ayuda
	5	Necesita ayuda. Realiza sin ayuda más de la mitad de estas tareas en un tiempo razonable
Arreglo	0	Dependiente. Necesita ayuda para las mismas
	5	Independiente. Realiza todas las actividades personales sin ayuda alguna; los complementos necesarios pueden ser provistos por alguna persona
Deposición	0	Dependiente. Necesita alguna ayuda
	10	Continente. No presenta episodios de incontinencia
	5	Accidente ocasional. Menos de una vez por semana o necesita ayuda para colocar enemas o supositorios
Micción	0	Incontinente. Más de un episodio semanal
	10	Continente. No presenta episodios. Capaz de utilizar cualquier dispositivo por si solo (botella, sonda, orinal, etc.)
	5	Accidente ocasional. Presenta un máximo de un episodio en 24 horas o requiere ayuda para la manipulación de sondas o de otros dispositivos
Ir al retrete	0	Incontinente. Más de un episodio en 24 horas
	10	Independiente. Entra y sale solo y no necesita ayuda alguna por parte de otra persona
	5	Necesita ayuda. Capaz de manejarse con una pequeña ayuda; es capaz de usar el cuarto de baño. Puede limpiarse solo
Transferencia (traslado cama/sillón)	0	Dependiente. Incapaz de acceder a él o de utilizarlo sin ayuda mayor
	15	Independiente. No requiere ayuda para sentarse o levantarse de una silla ni para entrar o salir de la cama
	10	Mínima ayuda. Incluye una supervisión o una pequeña ayuda física
	5	Gran ayuda. Precisa ayuda de una persona fuerte o entrenada
Deambulación	0	Dependiente. Necesita una grúa o el alzamiento por dos personas. Es incapaz de permanecer sentado
	15	Independiente. Puede andar 50 metros o su equivalente en casa sin ayuda supervisión. Puede utilizar cualquier ayuda mecánica excepto un andador. Si utiliza una prótesis, puede ponérsela y quitársela solo
	10	Necesita ayuda. Necesita supervisión o una pequeña ayuda física por parte de otra persona o utiliza andador
Subir y bajar escaleras	5	Independiente en silla de ruedas. No requiere ayuda ni supervisión
	10	Independiente. Capaz de subir y bajar un piso sin ayuda ni supervisión de otra persona
	5	Necesita ayuda. Necesita ayuda o supervisión
	0	Dependiente. Es incapaz

dios del IB previo evaluados por los médicos residentes que eran de 88,2 (21).

No obstante, en 40 pacientes (48%) existía alguna diferencia en el valor medio del IB. Al evaluar los pacientes en que existía diferencia en los valores medios del IB éstos eran pacientes de más edad (76,2 frente a 69,2;  $p < 0,005$ ), pero no había diferencias significativas respecto al género (65% mujeres frente a 45%;  $p = 0,6$ ) ni en la comorbilidad (2,8 frente a 2,1;  $p = 0,47$ ). Tampoco existían diferencias significativas según a quién se realizase la entrevista, si al paciente o al cuidador ( $p = 0,06$ ).

La tabla 1 muestra los resultados según los distintos apartados del IB. Destaca que el acuerdo interobservador era bajo ( $\kappa < 0,4$ ) en los apartados de comer, arreglo y trasferencias y solamente aceptable en el resto ( $\kappa$  entre 0,40-0,75).

## Discusión

El IB original y sus versiones fueron diseñadas para ser puntuadas por un observador entrenado<sup>2,7</sup>, incluso sin necesidad de que fuese un clínico, aunque las concordancias son mayores con observadores clínicos<sup>11</sup>. En nuestro estudio observamos que aunque las medias de la puntuación total son semejantes existen diferencias en el 48% de los valores medios de los pacientes evaluados; ello se debería a que en algunos apartados el médico en aprendizaje puntuaría al alza y en otros a la baja, y aunque la concordancia en la puntuación global es aceptable, ésta es baja en varios apartados, por lo que lo realmente valorable es la valoración apartado por apartado.

Se analizaron algunas posibles causas de discrepancias en los apartados en los que había más diferencias. En el apartado de «comer» la media de la enfermera era menor que la del médico; esto parecía ser debido a que en ocasiones el médico puntuaba con 10 puntos cuando en realidad el paciente necesitaba ayuda para cortar o extender mantequilla en el pan (serían 5 puntos). En el apartado del «arreglo» las diferencias podrían ser debidas a que en pacientes con dificultad para el afeitado o para lavarse la boca se puntuaba con 5 puntos debido a que no se preguntaba especí-

ficamente (serían 0 puntos) y en ocasiones se confundían con algunos aspectos del baño y se puntuaban de menos. Finalmente en el apartado de las «transferencias» se puntuaba por parte del médico residente ocasionalmente de menos; por ejemplo, cuando el paciente precisaba mínimas ayudas para sentarse en una silla se puntuaban 5 puntos (serían 10 puntos) y de más como, por ejemplo, cuando el paciente era incapaz de cambiar una silla de sitio que se puntuaba como 15 puntos (serían 10 puntos).

En el resto de apartados existía mejor correlación; no obstante, se apreciaron discrepancias ocasionales en los apartados de vestirse en que el médico en formación puntuaba 10 puntos cuando el paciente ocasionalmente necesitaba ayuda para ponerse las medias o atarse los zapatos, y en los apartados de caminar y subir o bajar escaleras en que en ocasiones se puntuaba lo que el paciente consideraba que podría desarrollar pero que hacía meses que no hacía o que normalmente no hacía pero que sí ocasionalmente (subir escaleras el día en que no funciona el ascensor). En el baño en ocasiones el médico no penalizaba la necesaria supervisión. En el apartado del retrete se penalizó incorrectamente algún paciente por usar cuña, aunque era autónomo para su uso, y en los apartados de incontinencia el principal problema fue en la correcta interpretación del término ocasional.

En conclusión creemos que la baja reproductividad interobservador en algunos apartados del IB entre enfermeras clínicas expertas en geriatría y médicos de familia en su período de formación en geriatría aconseja que es recomendable un período de aprendizaje para mejorar su aplicación.

## BIBLIOGRAFÍA

1. Valderrama Gama E, Pérez del Molinero Martín J. Una visión crítica de las escalas de valoración funcional traducidas al castellano. Rev Esp Geriatr Gerontol. 1997;32:297-306.
2. Mahoney FI, Barthel DW. Functional evaluation: the Barthel index. A simple index of independence useful in scoring improvement in the rehabilitation of the chronically ill. Md State Med J. 1965;14:61-5.
3. Cid Ruzafa J, Damián Moreno J. Valoración de la discapacidad física: el índice de Barthel. Rev Esp Salud Pública. 1997;2:127-37.
4. Batzán JJ, Pérez del Molino J, Alarcón T, San Cristóbal E, Izquierdo G, Manzarebitz J. Índice de Barthel: instrumento válido para la valoración funcional de pacientes con enfermedad cerebrovascular. Rev Esp Geriatr Gerontol. 1993;28:32-40.
5. Valverde MD, Flórez M, Sánchez I. Escala de actividades de la vida diaria. Rehabilitación. 1994;28:377-88.
6. Pfeiffer E. A short portable mental status questionnaire for the assessment of organic brain deficit in elderly patients. J Am Geriatr Soc. 1975; 23:433-41
7. Gresham GE, Phillips TF, Labi ML. ADL status in stroke: relative merits of three standard indexes. Arch Phys Med Rehabil. 1980;61:335-8.
8. Granger CV, Dewis LS, Peters NC, Sherwood CC, Barrett JE. Stroke rehabilitation analysis of repeated Barthel Index measures. Arch Phys Med Rehabil. 1979;60:14-7.
9. Collin C, Wade DT, Davis S, Horne V. The Barthel ADL Index: a reliability study. Int Disabil Stud 1988;10:61-3.
10. McDowell I, Newell C. Measuring health: a guide to rating scales and questionnaires. 2<sup>nd</sup> ed. Oxford: Oxford University Press; 1996.
11. Richards SH, Peters TJ, Coast J, Gunnell DJ, Daelow MA, Pounson J. Inter-rater reliability of the Barthel ADL index: how does a researcher compare to a nurse? Clin Rehabil. 2000;14:72-8.
12. Korner-Bitensky N, Wood-Dauphinee S. Barthel index information elicited over the telephone. Is it reliable? Am J Phys Med Rehabil. 1995;74:9-18.
13. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. J Chron Dis. 1987;40:378-83.
14. Dixon WJ, Brown MB, Engelmann L, Jennrich RI. BMDP statistical software manual. Berkeley: University of California Press, 1990.

**TABLA 1**  
**Concordancias interobservador (prueba kappa de Cohen) en los distintos apartados del índice de Barthel previo al ingreso**

	Evaluación enfermera	Evaluación médico	Valor kappa
Comer	9,4 (1,8)	9,8 (0,9)	0,327
Lavado (baño)	3 (2,4)	3,3 (2,4)	0,683
Vestido	8,6 (2,2)	8,7 (2,4)	0,571
Arreglo	4,6 (1,4)	4,6 (1,9)	0,395
Deposición	9,2 (2,2)	9,1 (2,3)	0,698
Micción	8,2 (3)	7,7 (3,2)	0,589
Ir al retrete	9,3 (2,1)	9,2 (2,2)	0,636
Trasferencias	14 (3,2)	13,5 (3)	0,268
Deambulación	14,2 (3,2)	13,2 (3,7)	0,576
Escaleras	7,9 (3,9)	7,5 (3,8)	0,403