

# Validación de un protocolo clínico para la detección de demencia en ámbito poblacional

C. Villanueva-Iza <sup>a</sup>, F. Bermejo-Pareja <sup>a</sup>, A. Berbel-García <sup>b</sup>, R. Trincado-Soriano <sup>a</sup>,  
J. Rivera-Navarro <sup>a</sup>

## VALIDATION OF A CLINICAL PROTOCOL FOR THE DETECTION OF DEMENTIA IN THE POPULATION

**Summary.** Aims. To analyse the validity of a set of neuropsychological and functional tests, and to study their value in detecting and diagnosing dementia through a pilot study. Patients and methods. A total of 131 subjects (101 controls and 30 with dementia) were evaluated using a comprehensive neuropsychological and functional battery. Validity analyses were conducted using ROC curves in accordance with the definitions of diagnostic test validation. Finally, a discriminant analysis was performed with the tests that showed greater diagnostic validity in the study of the ROC curves. Results. The case and control groups were not significantly different as regards age, sex and level of schooling. The ROC curves analyses showed the following to be the tests with the highest diagnostic validity: the MMSE, delayed recall of a short story, delayed recall of six pictures, the Spanish version of the S-IQCODE (shortened) and Pfeffer's FAQ. The discriminant analysis evidenced the fact that the joint utilisation of all the foregoing tests, except delayed recall of six pictures, classified 96.55% of our sample correctly. Conclusions. By combining direct cognitive evaluation of the subject and functional performance evaluated by a trustworthy informer, the vast majority of participants in the pilot study were correctly classified into patients with and without dementia. The high diagnostic validity of four relatively short tests lends support to their use in broader clinical or population studies. [REV NEUROL 2003; 36: 1121-6]

**Key words.** Dementia. Diagnosis. Neuropsychological and functional protocol. Validation.

## INTRODUCCIÓN

Los intervalos de prevalencia de demencia en la población anciana, mayor de 65 años, oscilan, según los estudios poblacionales, entre un 5 y un 10% en diversos países [1,2], y entre un 6 y un 14% en España [3]. Estas diferencias en las tasas de prevalencia se han atribuido a diversas causas, como el uso de diferentes criterios [4-7] e instrumentos diagnósticos [8-13] de demencia, así como a la dificultad diagnóstica que suponen los estadios iniciales de demencia; las diferencias en las características demográficas de las muestras poblacionales estudiadas [14-17] también tienen una indudable influencia.

Los test para el cribado se utilizan de manera generalizada en la práctica clínica y en estudios poblacionales. Estos instrumentos deberían ser fiables, con buena validez diagnóstica, de fácil aplicación y valoración y coste reducido. En la práctica, no hay un único instrumento que en sí mismo alcance este perfil ideal. Esto supone, con frecuencia, la utilización de varios instrumentos conjuntamente. En cualquier caso, en la actualidad, y tanto en la clínica como en la investigación poblacional, todavía no existe un claro acuerdo y una homogeneización en cuanto a las pruebas y cuestionarios que deben administrarse para el cribado y diagnóstico de la demencia. El propósito de nuestro estudio es analizar la validez de un conjunto de pruebas neuropsicológicas y cuestionarios para el informador en la detección de demencia en la población. La aplicación simultánea de dos fuentes de datos, una directa, a través del examen mental del paciente, y otra indirecta,

a través de un informador próximo y válido, incrementarían la información acerca de la existencia o no del deterioro cognitivo del sujeto, de su evolución en el tiempo y de su interferencia con las actividades sociolaborales, aspectos, todos ellos, centrales para el diagnóstico de demencia [5, 18-20]. No utilizar ambos tipos de información diferentes—directa e indirecta— puede ser la causa de que los test de cribado fracasen en la detección de demencia en los casos muy leves [18,21].

## PACIENTES Y MÉTODOS

### Población y métodos

#### Población

Se evaluaron 131 sujetos, de los cuales 101 eran controles y 30 eran casos con diagnóstico de demencia. Estos últimos procedían de la consulta de Neurología del Hospital Universitario 12 de Octubre, y se diagnosticaron de demencia con el acuerdo de dos neurólogos (para este estudio), ciegos al protocolo objeto de estudio y de acuerdo con los criterios DSM-III-R [5]. Se reclutaron, sucesivamente, de dicha consulta, desde junio de 1996 a junio del 1997, hasta alcanzar 30 casos. Los casos correspondieron a enfermos con el diagnóstico de demencia de intensidad leve-moderada de ámbito hospitalario con etiología degenerativa primaria (27 posible/probable Alzheimer, dos Lewy, uno frontotemporal), que se reevaluaron para este estudio. Los 101 sujetos del grupo control se eligieron de manera aleatoria de una muestra poblacional [6], con el diagnóstico de sujetos cognitivamente normales. Todos los sujetos eran mayores de 65 años y se estratificaron por edad, sexo y nivel cultural con los pacientes con demencia (Tabla I); 70 de estos sujetos pertenecían al área urbana (35 de Getafe, y 35 del distrito de Salamanca, ambos de Madrid) y los otros 31 procedían de un área rural (Arévalo, Ávila).

#### Diseño

El estudio se realizó mediante una comparación de dos grupos, previamente ajustados en variables demográficas y educativas. Se compararon los casos (serie clínica) con el grupo control (serie poblacional).

#### Procedimiento

Todos los casos se reclutaron de la consulta hospitalaria de Neurología del HU 12 de Octubre. Posteriormente, se evaluaron con la extensa batería objeto de este estudio, que se describirá a continuación. Realizaron la evaluación un psicólogo y un neurólogo, distintos de los que realizaron el diagnóstico clínico.

Recibido: 27.02.03. Aceptado tras revisión externa sin modificaciones: 31.03.03.

<sup>a</sup> Servicio de Neurología. Hospital Doce de Octubre. Madrid. <sup>b</sup> Sección de Neurología. Hospital de Alcázar de San Juan. Alcázar de San Juan. Ciudad Real, España.

Correspondencia: Dra. Clara Villanueva Iza. Cáceres, 6, 3.º C. E-28045 Madrid. E-mail: cvillaiza@telefonica.net

Agradecimientos. A Laboratorios Novartis, SA, por su colaboración y apoyo económico a un becario.

Estudio financiado con ayuda de la beca FIS00/0011-01.

© 2003, REVISTA DENEUROLOGÍA

Los sujetos pertenecientes al grupo control se eligieron aleatoriamente de la base del estudio NEDICES, y cumplimentaron un cuestionario estandarizado semejante al del primer corte [6] de este estudio, correspondiente a información demográfica, salud general –detección de las principales enfermedades crónicas del anciano y test de cribado de enfermedades neurológicas– y calidad de vida, además del mismo protocolo neuropsicológico que los casos. Esta evaluación la realizaron entrevistadores legos entrenados para tal fin.

#### Variables

Se recogieron las siguientes variables de los sujetos: sexo, edad (expresada en años) y nivel educativo –a los sujetos se les agrupó en analfabetos, con estudios primarios, con bachillerato y con estudios universitarios medios o superiores, y en todos los casos se recogió si eran completos o no–. Además, se recogieron las puntuaciones directas en las siguientes pruebas:

- *MMSE (del inglés, mini-mental state evaluation) modificada, versión de 37 puntos* [22]. Se utilizó la versión en español de dicho test, modificada y ampliada. Incluye orientación temporoespacial, memoria, atención, cálculo, capacidad visuoespacial y lenguaje. Las modificaciones que presenta esta versión respecto al test habitual son tres: para la evaluación de la capacidad visuoespacial añade la realización de dos círculos entrelazados; en la ejecución de órdenes, se agrega una visual, que consiste en la imitación de una lámina que muestra a un sujeto que levanta los brazos, y, finalmente, para la atención, se incorpora la repetición de cinco dígitos (1-3-5-7-9) en orden inverso al de su presentación. La puntuación de este test oscila entre 0 y 37 puntos.
- *Recuerdo inmediato y diferido de una historia (RIH/RDH) –memoria verbal de textos–* [23]. La historia consta de seis ideas, y se evalúa el recuerdo inmediato y el diferido a los 30 minutos de su presentación. El intervalo de puntuaciones varía de 0 a 6 puntos. Este test se inspira en la memoria de textos, de la escala de memoria revisada de Weschler.
- *Recuerdo inmediato y diferido (RIL/RDL) de láminas SEN-FIS* [24]. Se presentan seis láminas que corresponden a objetos comunes y se solicita al sujeto el recuerdo inmediato y, posteriormente, el recuerdo diferido a los 5 minutos de su presentación. La puntuación varía entre 0 y 6. Las láminas proceden del test Barcelona.
- *Fluidez verbal categorial (FVC)* [25]. Se recoge el número de elementos de una categoría –en esta caso, animales y frutas– que puede evocar un sujeto en un tiempo determinado –un minuto para cada una de ellas–. Es una tarea cognitiva compleja, en la que intervienen procesos lingüísticos, nemónicos y ejecutivos.
- *TMT (del inglés, trail making test)* [26-28]. Esta prueba se incluye en la batería neuropsicológica de Halstead-Reitan. Se realizan las partes A y B y se contabiliza el tiempo empleado para su realización. Se trata de un test para analizar el rastreo visual, la coordinación y rapidez visuomotora y la flexibilidad cognitiva. La parte A presenta una serie de círculos que contienen los números del 1-25, situados de forma aleatoria en el papel. El sujeto debe unir con una línea todos los dígitos en orden numérico lo más rápidamente posible. La parte B consta de números y letras distribuidos de forma aleatoria en el papel, y el sujeto debe unir en este caso, alternativamente, números y letras (número-letra-número-letra, etc.).
- *Recuerdo inmediato y diferido de cuatro palabras (RICP/RDCP)*. Se solicita el recuerdo inmediato y el recuerdo diferido, 15 minutos después de su presentación, y la puntuación varía de 0 a 4 puntos.
- *Test de acentuación de palabras (TAP)* [29]. Es una adaptación española del NART (del inglés, *national adult reading test*) de Nelson [30]. Se ha validado en castellano para determinar el nivel intelectual premórbido de los sujetos con sospecha de demencia. Consiste en la presentación de 30 palabras infrecuentes escritas sin tilde, y se evalúa su lectura con la acentuación correcta. La lectura de palabras aparece como resistente al deterioro cognitivo, y se mantiene intacta hasta fases avanzadas de la demencia [31]. La puntuación puede variar de 0 a 30 puntos.
- *Matrices progresivas de Raven (SPM)* [32]. Las series A y B de la escala de matrices progresivas de Raven se aplicaron a cada sujeto individualmente. Este test es una buena estimación de la inteligencia general. La puntuación, al emplear sólo dos series, varía entre 0 y 24.
- *Cuestionario del informador (SS-IQCODE)*. Es una versión española del cuestionario de Jorm y Korten [33]. Se ha validado en castellano y es una herramienta útil en la detección de demencia en los estadios precoces [34,35], aunque para este estudio se utilizó una versión acortada de seis ítems [36]. Consta de preguntas dirigidas al familiar o cuidador sobre los cambios ex-

**Tabla I.** Demografía de las dos series.

Variable	Dementes (n = 30)	No dementes (n = 101)	Significación
Edad media (DE)	75,06 (6,3)	77,57 (7)	NS
Sexo (V/M)	37/63%	46/54%	NS
Sin estudios (leen y escriben)	16,7%	24,8%	NS
Estudios primarios <sup>a</sup>	70%	59,4%	NS
Bachillerato <sup>a</sup>	10%	9,9%	NS
Estudios medios/superiores <sup>a</sup>	3,3%	5,9%	NS

DE: desviación estándar; NS: no significativo. <sup>a</sup>Se han agrupado estudios primarios completos e incompletos, bachillerato completo e incompleto y, por último, estudios medios y universitarios.

perimentados por el sujeto en los últimos 5 años en conductas de la vida diaria que implican procesos de memoria y otras capacidades intelectuales. La puntuación de cada ítem oscila de 1 a 5 (ha mejorado mucho, ha mejorado poco, apenas ha cambiado, ha empeorado un poco, ha empeorado mucho). La puntuación total es la suma de las puntuaciones obtenidas en todos los apartados. El intervalo de puntuación oscila, en este caso, entre 6 y 30.

- *Cuestionario de actividad funcional (FAQ de Pfeffer)* [37]. Es un cuestionario de actividades instrumentales de la vida diaria que se pasa al familiar o cuidador. Consta de 11 apartados. Cada uno se puntúa de 0 a 3 (0: normal; nunca lo hizo, pero podría hacerlo solo; 1: con dificultad, pero se maneja solo; nunca lo hizo y, si tuviera que hacerlo solo, tendría dificultad; 2: necesita ayuda, pero lo hace; 3: dependiente; no puede realizarlo). La puntuación total es la suma de las puntuaciones obtenidas en cada ítem, y oscila de 0 a 33 puntos; la discapacidad aumenta con la puntuación.
- *Índice de Katz (IK)* [38,39]. Cuestionario sobre las actividades básicas de la vida diaria, respondido por el familiar o cuidador. Consta de cinco apartados; cada ítem tiene tres respuestas posibles, y la barrera de dependencia para cada actividad depende del grado de ayuda que se necesita para realizarla. El intervalo de puntuación varía entre 5 y 15. Sólo se realizó en aquellos sujetos que presentaban una puntuación en la escala de Pfeffer mayor de 6.

#### Análisis estadístico

El análisis estadístico se realizó por medio del paquete estadístico SPSS [40] (SPSS Inc), 1989 de la Unidad de Investigación del HU 12 de Octubre. Se utilizó la *t* de Student para la comparación de medias, la prueba de ji al cuadrado ( $\chi^2$ ) con corrección de Fisher para variables categoriales, y test no paramétricos (Y de Mann-Whitney y test de Wilcoxon) para aquellas variables que no tenían una distribución normal o para variables categoriales que no debían analizarse mediante  $\chi^2$ .

Los análisis de validez se realizaron mediante las curvas ROC (del inglés, *receptor operative characteristics*), según las definiciones de validación de test diagnósticos [41]. Se establecieron los puntos de corte de las diferentes pruebas y cuestionarios y se tomó como más válido el punto en el que se obtenía una máxima concordancia.

Finalmente, se realizó un análisis discriminante entre los test que se determinaron como más válidos para conocer cuáles eran capaces de discriminar de forma más significativa entre sanos y enfermos. Se aplicó para ello el método Lambda de Wilks, y se tomaron los siguientes criterios de trabajo: tolerancia de 0,001; valor mínimo de *p* como criterio de entrada de las variables, 0,05, y valor máximo de *p* como criterio de salida, 0,1.

#### RESULTADOS

Se evaluaron 131 sujetos: 101 eran controles, y 30, dementes. La edad media (en años) del grupo con demencia fue de 75,1 (DE: 6,3), y la del grupo control, 77,6 (DE: 7). El 63% de los sujetos con demencia eran mujeres y el 54% de los controles. No se hallaron diferencias estadísticamente significativas entre ambos grupos (dementes y no dementes) en edad, sexo y nivel educativo (Tabla I).

**Tabla II.** Pruebas de la batería que no alcanzaron significación estadística.

	Dementes (n = 30)		No dementes (n = 101)		Significación
	Media	DE	Media	DE	
IK	6,1	1,64	7,4	2,17	0,08
TAP	13,6	8,8	16,3	9,3	0,18
Denominación de láminas SEN-FIS	5,66	0,6	5,79	0,58	0,31
RICP	3,53	1,1	3,91	0,42	0,07

DE: desviación estándar de la media; IK: índice de Katz; TAP: test de acentuación de palabras; RICP: recuerdo inmediato de cuatro palabras.

**Tabla III.** Pruebas de la batería que alcanzaron significación estadística.

	Dementes		No dementes		Significación
	Media	DE	Media	DE	
MMSE	21,8	4,2	31,26	4,12	< 0,0001
RIL	3,1	1,06	5,03	1,15	< 0,0001
RDL	0,53	1,13	3,59	1,49	< 0,0001
RIH	2,76	1,33	4,45	1,46	< 0,0001
RDH	0,33	1,51	3,96	1,94	< 0,0001
RDCP	0,2	0,55	1,02	1,21	< 0,0001
FVC animales	7,46	2,66	13,6	4,72	< 0,0001
FVC frutas	6,16	2,29	10,3	3,25	< 0,0001
SPM-A	4,93	3,11	8,27	2,15	< 0,0001
SPM-B	3,7	1,8	5,4	2,5	< 0,001
SPMA+B	8,56	4,46	13,71	4,11	< 0,0001
TMT-A	2,7	4,04	0,84	2,56	< 0,05
TMT-B	11,6	4,54	2,72	3,71	< 0,0001
SS-IQCODE	25,4	2,89	19,7	3,10	< 0,0001
FAQ	15,06	6,4	1,48	3,73	< 0,0001

DE: desviación estándar de la media; RIL: recuerdo inmediato de láminas; RDL: recuerdo diferido de láminas; RIH: recuerdo inmediato de la historia; RDH: recuerdo diferido de la historia; RDCP: recuerdo diferido de cuatro palabras; FVC: fluidez verbal categorial; SPM: matrices progresivas de Raven serie A, B y suma de las dos series. TMT: *trail making test*; SS-IQCODE: cuestionario abreviado del informador; FAQ: cuestionario de actividad funcional de Pfeffer.

**Tabla IV.** Áreas bajo la curva de las pruebas seleccionadas como más válidas.

	Áreas	IC 95%
FAQ	0,97	0,94-0,99
MMSE	0,94	0,91-0,98
RDL	0,93	0,86-0,99
RDH	0,91	0,85-0,97
SS-IQCODE	0,89	0,82-0,97

FAQ: cuestionario de actividad funcional de Pfeffer; RDL: recuerdo diferido de láminas; RDH: recuerdo diferido de historia; SS-IQCODE: cuestionario abreviado del informador.

Tampoco se encontraron diferencias estadísticamente significativas entre los dos grupos de sujetos en algunas de las pruebas cognitivas y cuestionarios, como el IK, el TAP, denominación de las láminas SEN-FIS y el RICP, aunque en esta última prueba la significación estadística se acerca a  $p = 0,05$  (Tabla II).

Se encontraron diferencias estadísticamente significativas entre dementes y no dementes en el MMSE, RIL y RDL, RIH y RDH, RDCP, FVC, series A y B de las SPM, TMT, SS-IQCODE y FAQ de Pfeffer (Tabla III). La utilidad diagnóstica de cada una de las pruebas de la batería que habían demostrado una significación estadística se evaluó mediante la realización de curvas ROC. Tras ello, se eligieron aquellas que encerraban un área bajo la curva mayor y, por lo tanto, a criterio de los investigadores y en función de los resultados, cuáles eran más válidas (Tabla IV). Además, se determinó el punto de corte para cada una de las pruebas, de acuerdo con aquel en el que se alcanzaba una mejor relación entre la sensibilidad y la especificidad (Tabla V). Las pruebas determinadas como más válidas fueron FAQ, MMSE, RDL, RDH y SS-IQCODE.

En la figura se puede observar la comparación de las curvas ROC de estas cinco pruebas. Finalmente, se realizó un análisis discriminante entre aquellas pruebas que se habían determinado como más válidas mediante las curvas ROC. Los resultados indicaron que, con tan sólo cuatro de esas cinco pruebas, se clasifica adecuadamente al 96,55% de nuestra muestra. Estas pruebas fueron SS-IQCODE, MMSE, FAQ y RDH.

## DISCUSIÓN

En nuestro estudio se aplicó un conjunto de pruebas neuropsicológicas que incluían la evaluación de un amplio abanico de funciones cognitivas [23-32], combinadas con la información funcional obtenida de un allegado a través de las escalas [33-39], que pretendían la detección [42] de demencia poblacional.

No hubo diferencias estadísticamente significativas entre el grupo de dementes y controles no dementes en variables demográficas (edad, sexo y nivel educativo), lo que facilitó el contraste de ambos grupos en las variables del protocolo del estudio. Tampoco hubo diferencias significativas entre ambos grupos en las siguientes pruebas utilizadas: TAP, denominación de las láminas SEN-FIS, RICP y IK.

Los resultados obtenidos con el TAP, por un lado, apoyan que la capacidad de lecturano sufre deterioro hasta fases avanzadas de la demencia [29,31,43], y, por otro, confirman la ausencia de diferencias en la capacidad intelectual premórbida entre controles y enfermos en nuestro estudio. Que no hubiera diferencias entre dementes y controles en el rendimiento en la denominación de láminas se debe, verosíblemente, a que el número de láminas empleadas se redujeron, y también, a que las imágenes presentadas son sencillas y familiares a los sujetos; por ello, parece razonable que no aparezcan alteraciones claras en esa prueba en las fases leve o moderada de la demencia, sino en fases más avanzadas, o si existiese un trastorno del lenguaje o perceptivo grave y precoz. La ausencia de diferencias entre ambos grupos de sujetos en el RICP podría explicarse de forma similar. Sólo en trastornos graves y precoces del lenguaje, o en fases muy avanzadas de la demencia, donde la memoria de fijación se encuentra muy alterada, el rendimiento en la prueba se vería presumiblemente afectado. Por último, es lógico que el IK [38,39], que evalúa actividades básicas de la vida diaria, no mostrara diferencias llamativas intergrupo, ya que no se realizó a todos los sujetos, sino sólo a aquellos que tenían una puntuación en el FAQ [37] mayor de 6; es decir, con un deterioro llamativo en tales actividades, ya se asocia o no a demencia.

El resto de las pruebas neuropsicológicas y los cuestionarios aplicados al informador demostraron diferencias con una alta significación estadística ( $p < 0,001$  y  $p < 0,0001$ ). Al analizar mediante curvas ROC cuáles de ellas tenían una mayor validez diagnóstica, se observó que las cinco mejores pruebas eran el MMSE de 37 apartados [23], el RDH, el RDL [24], el SS-IQCODE [36] y el FAQ [37].

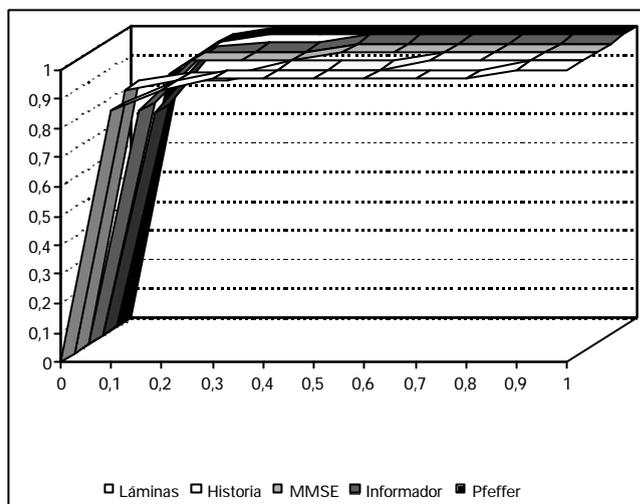
El S-IQCODE en su versión de 26 y 17 apartados, se ha utilizado en otros estudios, tanto con muestras clínicas como poblacionales [33-35,44], y ha demostrado su validez diagnóstica. La versión abreviada de 6 apartados [36] empleada con nuestra muestra reflejó una sensibilidad del 93,3%, una especificidad del 84,2%, un valor predictivo positivo del 75,7%, un valor predictivo negativo del 94,6% y una concordancia del 0,87 –valores dentro del intervalo de los obtenidos en los estudios con las versiones más largas–.

Las escalas instrumentales, como el FAQ [37,45], son muy sensibles a los cambios en las demencias leves. Nuestros datos reflejan una buena validez diagnóstica, con una sensibilidad del 93,3%, una especificidad del 90,1%, un valor predictivo positivo del 73,3%, un valor predictivo negativo del 97,8% y una concordancia de 0,91. Lo cual, como en el caso del SS-IQCODE, nos llevaría a plantear su posible inclusión en futuros protocolos para la detección de demencia y deterioro cognitivo. El empleo conjunto de ambos cuestionarios aporta información externa al sujeto sobre la existencia o no de deterioro cognitivo y su impacto en la vida cotidiana de dicho sujeto; además, no sólo se evalúa el deterioro en el momento de la administración de las pruebas, sino también la evolución y magnitud de ese deterioro [20,44,46]; ello resulta fundamental, especialmente en estadios precoces, donde el rendimiento cognitivo puede permanecer todavía muy levemente alterado [47].

Mediante el recuerdo diferido de una historia corta [RDH], en nuestra muestra obtuvimos una sensibilidad del 90%, una especificidad del 89,1%, un valor predictivo positivo del 76,1%, un valor predictivo negativo del 96,8% y una concordancia de 0,89, todo lo cual creemos que refleja unos buenos valores para plantear su utilización como herramienta de cribado. No tenemos conocimiento hasta la fecha de su utilización con muestras poblacionales.

El MMSE se ha utilizado como herramienta de cribado tanto en estudios poblacionales como clínicos. El punto de corte de 23-24 se ha asumido como el punto de corte de mayor validez diagnóstica y se emplea en la gran mayoría de los estudios en la bibliografía anglosajona [48]. No obstante, es importante destacar que todavía persiste el debate sobre la necesidad de adaptar o no el punto de corte a las características sociodemográficas de los sujetos en los que se aplica este test y al objetivo de la investigación [49,50]. Nosotros empleamos una modificación y ampliación del MMSE de 37 ítems [23], con un punto de corte menor de 25; se obtuvo una sensibilidad del 80%, una especificidad del 92,1%, un valor predictivo positivo del 75%, un valor predictivo negativo del 93,9% y una concordancia del 0,89, valores dentro de los descritos. Pensamos que los resultados obtenidos son lógicos, ya que, al igual que otras versiones del MMSE, mide las mismas funciones del estado mental.

Finalmente, el análisis discriminante llevado a cabo redujo a cuatro las pruebas que clasifican correctamente a la población objeto de este estudio. El análisis eliminó el RDL, posiblemente porque la información aportada por esta prueba resultaba redundante, si tenemos en cuenta que el RDH también valora la memoria. En nuestro estudio, con los cuatro test reseñados se clasificó adecuadamente al 96,5% de la muestra.



**Figura.** Gráfica comparativa de las curvas ROC obtenidas para aquellas pruebas que resultaron ser más válidas. Eje X: 1 – especificidad; eje Y: sensibilidad.

**Tabla V.** Validación diagnóstica mediante curvas ROC. Puntos de corte determinados como más válidos para las cinco pruebas.

Prueba	S	E	VPP	VPN	CONC	LR+	PC
RDH	90	89,1	76,1	96,8	0,89	8,1	<1
FAQ	93,3	90,1	73,7	97,8	0,91	9,3	35
RDL	86,7	90,1	72,2	95,8	0,89	10	<2
MMSE	80	92,1	75	93,9	0,89	10	<25
SS-IQCODE	93,3	84,2	75,7	94,6	0,87	5,8	>22

S: Sensibilidad (%); E: especificidad (%); VPP: valor predictivo positivo (%); VPN: valor predictivo negativo (%); CONC: concordancia; LR+: *likelihood ratio* positivo; PC: punto de corte estimado; RDH: recuerdo diferido de la historia; RDL: recuerdo diferido de láminas; SS-IQCODE: cuestionario del informador abreviado; FAQ: cuestionario de actividad funcional de Pfeffer.

Como conclusión, pensamos que el empleo conjunto de tests cognitivos y cuestionarios estructurados para el informador es un instrumento válido para la detección de demencia en muestras poblacionales, ya que reúne datos del deterioro cognitivo puntual del sujeto, la evolución del mismo y el impacto social que provoca. Recordemos que en nuestro estudio, con el empleo de FAQ + MMSE + RDH + SS-IQCODE se clasificaba adecuadamente al 96,5% de la muestra, un porcentaje muy elevado.

El pequeño tamaño de la muestra de casos, propia de un estudio piloto, y la uniformidad del diagnóstico de los casos de demencia limita la validez de estos resultados; en cambio, la ausencia de demencia grave hace suponer que en una muestra más amplia el conjunto de pruebas alcanzaría una validez también notable. Creemos que los resultados permiten explorar este conjunto de pruebas neuropsicológicas y funcionales en muestras clínicas o poblacionales más amplias.

#### BIBLIOGRAFÍA

- Henderson AS. Demencia. Epidemiología de los trastornos mentales y de los problemas psicosociales. Madrid: Meditor; 1996.
- Bermejo F. Epidemiología de la enfermedad de Alzheimer; situación al comienzo de una década. European Journal of Gerontology 1992; (Suppl): 5-18.
- López-Pousa S, Vilalta J, Llinás J. Epidemiología de las demencias en España. Revista de Gerontología 1995; (Suppl 1): 28-33.
- American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 3 ed. Washington: American Psychiatric Association; 1980.

5. American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 3 ed revised. Washington: American Psychiatric Association; 1987.
6. Bermejo F, Gabriel R, Vega S, Morales JM, Rocca WA, Anderson DW. Problems issues with door-to-door, to-fase surveys: an illustration from Central Spain. *Neuroepidemiology* 2001; 20: 225-31.
7. MacKahn G, Drachman D, Folstein M, Katzman R, Price D, Standlar EM. Clinical diagnosis of Alzheimer's disease: Reports of the NINCDS-ADRDA work group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology* 1984; 34: 939-44.
8. Folstein MF, Folstein SE, McHugh PR. Mini-mental state. A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* 1975; 12: 397-408.
9. Lobo A, Ezquerro J, Gómez FB, Sala JM, Seva-Díaz A. El miniexamen cognoscitivo. Un test sencillo, práctico, para detectar alteraciones intelectivas en pacientes médicos. *Actas Luso Esp Neurol Psiquiatr Cienc Afines* 1979; 3: 189-202.
10. Roth M, Tym E, Mountjoy CQ, Huppert FA, Hendrie H, Verma S, et al. CAMDEX: a standardized instrument for the diagnosis of mental disorders in the elderly with special reference to the early detection of dementia. *Br J Psychiatry* 1986; 149: 698-709.
11. Llinás J, Vilalta J, López-Pousa S. CAMDEX: adaptación y validación española. Barcelona: Ancora; 1991.
12. Hodkinson M. Abreviated mental state. *Age Ageing* 1972; 1: 233-8.
13. Copeland JR, Dewey ME, Griffith-Jones M. Computerized psychiatric diagnosis system and case nomenclature for elderly subjects: GMS and AGE-CAT. *Psychol Med* 1986; 16: 89-99.
14. Bermejo F. Epidemiología de la demencia senil en España. Situación actual y perspectivas futuras. In Tolosa E, Bermejo F, Boller F, eds. *Demencia senil*. Barcelona: Springer-Verlag Ibérica; 1991. p. 1-15.
15. Coria F, Gómez de Caso JA, Mínguez L, Rodríguez-Artalejo F, Clavería LE. Prevalence of age-associated memory impairment and dementia in a rural community. *J Neurol Neurosurg Psychiatry* 1993; 56: 973-6.
16. Lobo A, Dewey M, Copeland J, Días JL, Saz P. The prevalence of dementia among elderly people living in Zaragoza and Liverpool. *Psychol Med* 1992; 22: 239-43.
17. López-Pousa S, Vilalta J, Llinás. Prevalence of dementia in old age in Girona, Spain. *Neuroepidemiology* 1991; 10: 99.
18. Morris JC, Fulling K. Early Alzheimer's disease. Diagnostic considerations. *Arch Neurol* 1988; 45: 345-9.
19. Eslinger PJ, Damasio AR, Benton AL, Allen M. Neuropsychologic detection of abnormal mental decline in older persons. *JAMA* 1985; 253: 670-4.
20. Applegate WB, Blass JP, Williams TF. Instruments for the function assessment of older patients. *N Engl J Med* 1990; 322: 1207-14.
21. Nelson A, Fogel B, Faust D. Bedside cognitive screening instruments. A critical assessment. *J Nerv Ment Dis* 1986; 174: 73-83.
22. Bermejo F, Díaz J, Olazarán J. Prevalencia de demencia en una muestra poblacional de ancianos en medio urbano. Análisis preliminar. *Neurología* 1995; 10: 418.
23. Wechsler D. Manual for the Memory Scale-Revised (WMS-R). San Antonio, TX: The Psychological Corporation; 1985.
24. Grupo de estudio de demencias. Protocolo de estudio multicéntrico de casos incidentes de demencia en España. Madrid: SEN. Grupo de Estudio de Demencias; 1983.
25. Lezak MD. Verbal functions and language skills. In Lezak MD, eds. *Neuropsychological assessment*. New York: Oxford University Press; 1995. p. 523-55.
26. Lezak MD. Batteries for assessment of brain damage. In Lezak MD, eds. *Neuropsychological assessment*. New York: Oxford University Press; 1995. p. 559-98.
27. Lezak MD. Constructions. In Lezak MD, eds. *Neuropsychological assessment*. New York: Oxford University Press; 1995. p. 686-735.
28. Pascual-Millán LF. Las praxias constructivas. In Peña-Casanova J, ed. *La exploración neuropsicológica*. Barcelona: MCR; 1988. p. 125-38.
29. Del Ser T, González-Montalvo JI, Martínez-Espinosa S, Delgado-Villalpalos C, Bermejo F. Estimation of premorbid intelligence in Spanish people with the Word Accentuation Test and its application to the diagnosis of dementia. *Brain Cogn* 1997; 33: 343-56.
30. Nelson HE, O'Connell A. Dementia: the estimation of premorbid intelligence levels using the new adults reading test. *Cortex* 1987; 14: 234-44.
31. O'Carroll RE, Baikie EM, Wittick JE. Does the national adult reading test hold in dementia? *Br J Clin Psychol* 1987; 26: 315-6.
32. Raven JC, Court JH, Raven J. *Matrices progresivas*. Manual en español. Madrid: TEA, Publicaciones de Psicología Aplicada; 1995.
33. Jorm AF, Korten E. Assessment of cognitive decline in the elderly by informant interview. *Br J Psychiatry* 1988; 152: 209-13.
34. Morales JM, González-Montalvo JI, Bermejo F, Del Ser T. The screening of mild dementia with a shortened Spanish version of the informant Questionnaire on Cognitive Decline in the Elderly. *Alz Dis Assoc Disord* 1995; 9: 105-11.
35. Del Ser T, Morales JM, Barquero MS, Cantón R, Bermejo F. Application of a Spanish version test of the Informant Questionnaire on Cognitive Decline in the Elderly in the clinical assessment of dementia. *Alzheimer Dis Assoc Disord* 1997; 11: 3-8.
36. Morales-González JM. Validación del test del informador como test de screening de demencia en la población anciana [tesis doctoral]. Madrid: Universidad Complutense; 1994.
37. Pfeffer RI, Kurosaki TT, Harrah CH, Chance JM, Filos S. Measurement of functional activities in older adults in the community. *J Gerontol* 1982; 37: 323-9.
38. Katz S, Stroud MV. Functional assessment in geriatrics. A review of progress and directions. *J Am Geriatr Soc* 1989; 37: 267-71.
39. Katz S, Ford AB, Moskowitz RW, Jackson BS, Jaffe MW, Cleveland MA. Studies of illness in the aged. The index of ADL: a standardized measure of biological and psychosocial function. *JAMA* 1963; 185: 914-9.
40. SPSS Inc. *SPSS user's guide*. New York: McGraw-Hill; 1986.
41. Sackett DL, Haynes BR, Tugwell P. *Clinical epidemiology: a basic science for clinical medicine*. Boston: Little Brown; 1991.
42. Baldereschi M, Meneghini F, Quiroga P, Albala C, Mamo J, Gabriel R, et al. Cognitive versus functional screening for dementia across different countries: cross-cultural validation of the mini-mental state examination [MMSE] and of the Pfeffer functional activities questionnaire (FAQ) against the standardized clinical diagnosis of dementia. *Neurology* 1994; 44 (Suppl 2): A365.
43. Cummings JL, Houlihan JP, Hill MA. The pattern of reading deterioration in dementia of the Alzheimer type: Observations and implications. *Brain Language* 1986; 29: 315-23.
44. Jorm A, Scott R, Cullen JS, Mackinnon AJ. Performance of the informant questionnaire on cognitive decline in the elderly [IQCODE] as a screening of dementia. *Psychol Med* 1989; 21: 785-90.
45. Pfeffer RI, Kurosaki TT, Harrah CH, Bates D, Detels R, et al. A survey diagnostic tool for senile dementia. *Am J Epidemiol* 1981; 114: 115-27.
46. Huppert FA, Tym E. Clinical and neuropsychological assessment of dementia. *Br Med Bull* 1986; 42: 11-8.
47. Morris JC, McKell DW, Storandt M, Rubin EH, Price JL, Grant EA, et al. Very mild Alzheimer's disease: informant-based clinical, psychometric, and pathological distinctions from normal aging. *Neurology* 1991; 41: 469-78.
48. Tombaugh TN, McIntyre NJ. The mini-mental state examination: a comprehensive review. *J Am Geriatr Soc* 1992; 40: 922-35.
49. Bermejo F, Gómez-Isla T, Morales JM. El mini-mental state examination en la evaluación del deterioro cognitivo y la demencia. In Del Ser T, Peña J, eds. *Evaluación neuropsicológica y funcional de la demencia*. Barcelona: Prous; 1994. p. 93-107.
50. Bermejo F, Morales JM, Valerga C, Del Ser T, Artolazábal J, Gabriel R. Comparación entre dos versiones españolas abreviadas de evaluación del estado mental en el diagnóstico de demencia. Datos de un estudio en ancianos residentes en la comunidad. *Med Clin (Barc)* 1999; 112: 330-4.

**VALIDACIÓN DE UN PROTOCOLO CLÍNICO  
PARA LA DETECCIÓN DE DEMENCIA  
EN ÁMBITO POBLACIONAL**

**Resumen.** Objetivo. Analizar la validez de un conjunto de pruebas neuropsicológicas y funcionales y estudiar su utilidad en la detección y diagnóstico de la demencia mediante un estudio piloto. Pacientes y métodos. Se evaluaron 131 sujetos (101 controles y 30

**VALIDAÇÃO DE UM PROTOCOLO CLÍNICO  
PARA A DETECÇÃO DA DEMÊNCIA  
EM ÁMBITO POPULACIONAL**

**Resumo.** Objetivo. Analisar a validade de um conjunto de provas neuropsicológicas e funcionais, e estudar a sua utilidade na detecção e diagnóstico da demência através de um estudo piloto. Doentes e métodos. São avaliados 131 indivíduos (101 controles e 30 doen-

dementes) con una amplia batería neuropsicológica y funcional. Los análisis de validez se llevaron a cabo mediante la utilización de curvas ROC según las definiciones de validación de test diagnósticos. Finalmente, se realizó un análisis discriminante con las pruebas que mostraron mayor validez diagnóstica en el estudio de curvas ROC. Resultados. Los grupos de casos y controles no diferían significativamente en edad, sexo y nivel educativo. Los análisis de curvas ROC determinaron como pruebas con mayor validez diagnóstica las siguientes: MMSE, recuerdo diferido de una historia corta, recuerdo diferido de seis láminas, la versión española del S-IQCODE (reducida) y el FAQ de Pfeffer. El análisis discriminante reflejó que el empleo conjunto de las pruebas anteriores—todas, salvo el recuerdo diferido de seis láminas—, clasificaba al 96,55% de nuestra muestra correctamente. Conclusiones. La combinación de la evaluación cognitiva directa del sujeto y el rendimiento funcional evaluado por un informador fiable permite la clasificación correcta en dementes y no dementes de la gran mayoría de los participantes en el estudio piloto. La elevada validez diagnóstica de cuatro pruebas relativamente cortas sugiere su empleo en estudios más amplios, clínicos o poblacionales. [REV NEUROL 2003; 36: 1121-6]

**Palabras clave.** Demencia. Diagnóstico. Protocolo neuropsicológico y funcional. Validación.

tes), com uma ampla bateria neuropsicológica e funcional. As análises de validade foram efectuadas com a utilização de curvas ROC segundo as definições de validação de testes diagnósticos. Finalmente, realizou-se uma análise discriminativa com as provas que mostraram maior validade diagnóstica no estudo de curvas ROC. Resultados. Os grupos de casos e controles não diferiam significativamente por idade, sexo, e nível educativo. As análises de curvas ROC determinaram como provas com maior validade diagnóstica: o MMSE, recordação diferida de uma história breve, recordação diferida de seis estampas, a versão espanhola do S-IQCODE (reduzida) e o FAQ de Pfeffer. A análise discriminativa revelou que o emprego conjunto das provas prévias—todas, excepto a recordação diferida de seis estampas—, classificava correctamente 96,55% da nossa amostra. Conclusões. A combinação da avaliação cognitiva directa do indivíduo e o rendimento funcional avaliado por um informador fiável permite a classificação correcta em dementes e não dementes da grande maioria dos participantes no estudo piloto. A elevada validade diagnóstica de quatro provas relativamente curtas sugere a sua utilização em estudos clínicos ou populacionais mais amplos. [REV NEUROL 2003; 36: 1121-6]

**Palavras chave.** Demência. Diagnóstico. Protocolo neuropsicológico e funcional. Validação.