

VALIDEZ Y CONFIABILIDAD DEL MINIMENTAL STATE EXAMINATION (MMSE) Y DEL MMSE MODIFICADO (3MS) PARA EL DIAGNÓSTICO DE DEMENCIA EN MÉRIDA, VENEZUELA.

Clara Isabel Ramírez¹, Carlos Elí A. Moncada Rodríguez², Trino Baptista³

¹Unidad de Neurología, Facultad de Medicina, Universidad de Los Andes-Instituto Autónomo Hospital Universitario de Los Andes. ²Unidad de Embriología. Universidad de Los Andes. ³Unidad de Fisiología Humana, Universidad de Los Andes. E-mail: isabellarami66@hotmail.com

Resumen

Se evaluaron la validez y la confiabilidad del *Mini Mental State Examination* (MMSE) y de una versión ampliada del mismo: El *Minimental State Examination* modificado (3MS), diseñado para optimizar el diagnóstico de demencia. El estudio se hizo en el Instituto Autónomo Hospital Universitario de los Andes, Mérida, Venezuela. La muestra fue de 209 pacientes. Ambos instrumentos fueron semejantes en validez de contenido y de criterio concurrente, concordancia intra e inter observador y concordancia interna. La validez del constructo mostró tres factores para MMSE y cuatro para 3MS. La concordancia con el diagnóstico clínico (sano-demencia), expresada en unidades kappa, fue: 0.53 para 3MS y 0.39, para MMSE. Con tres categorías (sano-deterioro cognitivo leve- demencia) fue 0.50, para el 3MS y 0.27 para el MMSE. Sensibilidad: 3MS, 68%; MMSE, 42.0%. Especificidad: 3MS, 85.3%; MMSE, 96.3%. Valor predictivo positivo: 3MS, 39.4%; MMSE, 61.9%. Valor predictivo negativo: 92.2% y 95.0%, respectivamente. Validez discriminante: 3MS, 77.0%; MMSE, 70.3%. Área bajo la curva ROC: 3MS, 0.85; MMSE, 0.88. La edad y el nivel educativo influyeron significativamente en los resultados. Hubo dificultad para detectar deterioro cognitivo leve. Pacientes erróneamente diagnosticados sanos: 46.9% en el 3MS y 85.7% en el MMSE. Organizados los instrumentos en tres categorías, el resultado se mantuvo en el MMSE y disminuyó en el 3MS. Recomendamos disminuir el punto de corte del 3MS dicotómico, u organizarlo en tres categorías diagnósticas.

Palabras Clave: Demencia, pruebas diagnósticas, MMSE, 3MS, validez, confiabilidad

Abstract

Validity and reliability of the Mini-mental State Examination (MMSE) and the modified MMSE (3MS) for the diagnosis of dementia in Merida, Venezuela.

We assessed the validity and reliability of the Mini Mental State Examination (MMSE) and an its extended version: The modified Mini-mental State Examination (3MS), designed to optimize the diagnosis of dementia. The study was done at the Instituto Autónomo Hospital Universitario de los Andes, Mérida, Venezuela. The sample comprised 209 patients. Both instruments were similar in content validity and concurrent criterion, intra and inter-observer agreement and internal consistency. Construct validity for MMSE showed three factors and 4 for 3MS. The agreement with the clinical diagnosis (healthy-dementia), expressed in units kappa was: 0.53 for 3MS and 0.39 for MMSE. With three categories (healthy-mild cognitive impairment, dementia) was 0.50 for the 3MS and 0.27 for the MMSE. Sensitivity: 3MS, 68.0%; MMSE, 42.0%. Specificity: 3MS, 85.3%; MMSE, 96.3%. Positive predictive value: 3MS, 39.4%; MMSE, 61.9%. Negative predictive value: 92.2% and 95% respectively. Discriminant validity: 3MS, 77.0%; MMSE, 70.3%. Area under the ROC curve: 3MS, 0.85; MMSE, 0.88. The age and education level significantly influenced the results. There was difficulty in detecting mild cognitive impairment. Misdiagnosed patients healthy: 46.9% in the 3MS and 85.7% in the MMSE. Instruments organized into three categories, the result remained the MMSE and 3MS decreased in. We recommend reduce the 3MS cutoff dichotomous, or organized into three diagnostic categories.

Keywords: Dementia, diagnostic tests, MMSE, 3MS, validity, reliability.

INTRODUCCIÓN

Hasta hace algunos años la demencia se consideraba de poca importancia en el ejercicio práctico del médico clínico, pues se observaba en ocasiones como la “evolución normal en las personas de edad avanzada”. Sin embargo, la demencia se ha convertido en uno de los trastornos más importantes a medida que la expectativa de vida ha ido aumentando (Heber y Brayne 1995, Jorm et al. 1987). Actualmente se considera la demencia como un síndrome clínico de carácter orgánico, que produce un deterioro progresivo y global de las

funciones intelectuales. La demencia interfiere con la vida social o profesional del paciente, sin afectar, en general, el nivel de conciencia. El deterioro cognoscitivo puede ir desde una afectación leve casi imperceptible hasta un deterioro psíquico avanzado (Fischer 1951, Del Ser et al. 1993b).

El diagnóstico de la demencia tipo Enfermedad de Alzheimer (EA) inicial y leve se realiza en muy baja proporción. Apenas se identifica en la atención primaria, y su detección depende de la percepción de la familia sobre los cambios de hábitos y actitudes del enfermo.

El Examen Mental Abreviado de Folstein (MMSE), es uno de los instrumentos más ampliamente usados como prueba de evaluación del estado cognoscitivo y como prueba de tamizaje para la detección de demencia (Folstein et al. 1975, Kukull et al. 1994). Esta prueba ha demostrado validez y confiabilidad en numerosos estudios (Anthony et al. 1982).

Parte de la popularidad del MMSE puede ser atribuida a que es objetiva, breve, fácil de aplicar, y capaz de identificar a las personas que padecen estados avanzados de demencia (Folstein 1982).

Teng y Chang (1987) plantearon una modificación del MMSE para obtener el 3MS, que consistió en agregar cuatro ítems y extender el rango de puntuación. Estas modificaciones fueron diseñadas para mostrar una amplia variedad de funciones cognoscitivas, con el fin de cubrir un rango mayor de niveles de dificultad y de esta manera mejorar la confiabilidad y validez de la puntuación.

En el presente estudio se realizó la validación del 3MS que mantiene la brevedad de la prueba pero tiene modificaciones importantes como son: Extiende los rangos de puntuación del test hasta 100 puntos; incluye más pruebas para las habilidades cognitivas; y aumenta la confiabilidad y validez del puntaje. Estas modificaciones no han impedido que el MMSE se siga utilizando para la detección de la demencia, pero en vista de que se ha postulado que el 3MS puede ser más adecuado para la detección precoz de una proporción mayor de sujetos con demencia incipiente (Teng y Chang 1987), nos hemos propuesto realizar el estudio de validez y confiabilidad del 3MS.

METODOLOGÍA.

Con el fin de determinar las características operativas del 3MS, su validez y su confiabilidad en el diagnóstico de demencia de la población de pacientes mayores de 50 años, que acudieron a la consulta de las Unidades de Psiquiatría y Neurología del (I.A.H.U.L.A.), entre enero y junio del 2006, estimada en 487, se evaluó una muestra de 100 pacientes afectados de algún grado de deterioro cognitivo o demencia y 109, sin deterioro cognitivo.

Evaluación clínica de los sujetos del estudio principal:

Los pacientes que cumplieron con los criterios de inclusión, fueron evaluados con base en las exigencias del DSM-IV del NINCDS/ADRDA y la Escala de Ponderación de la Demencia (CDR), y ubicados en una de las categorías caso/no caso (American Psychiatric Association 1994, Tierney et al. 1988, López et al. 1990, Morris 1993).

Validez de contenido de los instrumentos: Se utilizó el coeficiente de validez de contenido (CVC)

desarrollado por Hernández Nieto en 2002. Para tal fin, seis jueces con experiencia extensa en el diagnóstico y tratamiento de la demencia evaluaron cada uno de los ítems de los dos instrumentos y lo ubicaron en una escala de tipo Likert, de acuerdo con la capacidad del ítem para valorar óptimamente la variable o dimensión que pretende explorar. Tanto en MMSE como en 3MS las dimensiones son: orientación, memoria, atención y cálculo, evocación, lenguaje y praxis. La escala Likert fue diseñada así: 1. Completamente en desacuerdo; 2. Moderadamente en desacuerdo; 3. Levemente en desacuerdo; 4. Levemente de acuerdo; 5. Moderadamente de acuerdo; 6. Completamente de acuerdo.

Validez de criterio: La validez de criterio concurrente se obtuvo correlacionando tanto las puntuaciones del MMSE como los del 3MS con la CDR (Hernández et al. 2003, Morris 1993).

Para realizar este análisis, la puntuación individual en cada instrumento se ubicó en una de tres categorías: Sin deterioro, deterioro leve y demencia.

Confiabilidad del MMSE y del 3MS: **A.** Se compararon las evaluaciones realizadas por cada entrevistador consigo mismo y con la de otro evaluador con el fin de cuantificar la asociación estadística. **B.** Índice de consistencia interna: Se aplicó el coeficiente de Cronbach (Hernández et al. 2003), el cual se consideró muy bueno por encima de 0.80; bueno de 0.70 a 0.80; medio de 0.45 a 0.60 y malo por debajo de 0.45. **C.** Para determinar la concordancia, se comparó el resultado obtenido en cada uno de los instrumentos con el diagnóstico clínico; la información obtenida de cada sujeto del estudio se codificó en una tabla tetracórica. Se consideró aceptable el grado de concordancia mayor o igual a 0.45 kappa (Ruiz 2002, Dawson y Trapp 2002, Hernández 2003).

Sensibilidad y especificidad: Se determinó la co-positividad y la co-negatividad con relación al diagnóstico clínico. Sensibilidad = $a/a+c$. Especificidad = $d/d+b$. Valor predictivo positivo = $a/a+b$. Valor predictivo negativo = $d/d+c$. Exactitud o confiabilidad discriminante = $(a+d)/a+b+c+d$. (Ruiz 2002).

Determinación de las curvas operativas para el receptor (COR): Se calcularon las curvas COR óptimas y sus correspondientes ABC para el MMSE y para el 3MS mediante el llamado método *no paramétrico* o *empírico* (López y Pita 1998).

Variables demográficas y desempeño en el MMSE y en el 3MS. Para este análisis inferencial se usó la tabla tetracórica y se utilizaron las pruebas de χ^2 y exacta de Fisher con su correspondiente razón de verosimilitud. Se calculó el riesgo relativo indirecto

(odds ratio) de la presencia de deterioro de acuerdo a las diversas categorías de las variables demográficas. Para ello, se utilizó la siguiente fórmula referida a la tabla tetracórica: Riesgo relativo indirecto = ad/bc.

Concordancia Global entre el EMA y el 3MS: Se examinó en que medida ambos instrumentos concordaron, en términos de unidades kappa, en la evaluación del estado mental.

Análisis estadístico: Los diversos análisis se realizaron con el programa SPSS (13.0). Los resultados se consideraron estadísticamente significativos cuando p fue < 0.05 .

RESULTADOS.

Características demográficas de la muestra:

De los 209 sujetos estudiados 41 (19.60%) fueron del sexo masculino, y 168 (80.40%) del femenino, 123 (58.9%) tenían entre 50 y 70 años de edad, y 86 (41.10 %) entre 71 y 86 años.

Sin deterioro cognitivo eran 109 personas, 49 tenían deterioro cognitivo leve y 51 presentaron demencia. Con primaria incompleta fueron 105 pacientes, Con secundaria incompleta y normalistas 73 y con educación superior 28. En tres pacientes no se obtuvo información.

Los coeficientes de variación proporcional empírica, obtenidos de las puntuaciones alcanzadas por los pacientes en el MMSE y en el 3MS de acuerdo con las variables demográficas y el nivel educativo, se presentan en las tabla 1.

Tabla 1. Coeficiente de variación proporcional empírica (CVPe) en los puntajes MMSE y 3MS de acuerdo con la edad y el nivel educativo.

Grupos	Categorías	N	CVPe %	
			MMSE	3MS
Sexo	Masculino	41	49,89	47,41
	Femenino	168	33,38	38,13
Edad	50 a 70 años	123	45,33	45,33
	71 a 86 años	86	39,79	39,79
Escolaridad	Primaria completa o incompleta	105	37,10	37,10
	Secundaria incompleta y normalistas	73	29,91	29,91
	Superior	28	41,38	41,38
Total		209	33,17	33,17

Escala de variabilidad: 0.00 –25, muy baja; 26 – 50, moderadamente baja; 51 – 75, moderadamente alta; 76 – 100, muy alta.

Validez de Contenido

El CVC total para el MMSE, alcanzó a 0.8906 y el corregido por concordancia aleatoria a 0.8903.

El CVC total para el 3MS, fue de 0.9575 y el corregido por concordancia aleatoria 0.9571. Como puede verse el CVC fue discretamente superior en el 3MS.

Validez de Criterio Concurrente.

MMSE, tabulado en tres categorías: demencia, deterioro leve y sin deterioro, y la escala CDR exhibieron una correlación de Kendall de -0.588 , $p = 0.000$, para $N = 209$.

La correlación de Kendall entre el 3MS (demencia, deterioro leve y sin deterioro) y CDR fue de: $r = -0.554$, $p = 0.000$, $N = 209$. En ambos instrumentos la correlación con CDR, fue alta e inversa.

Validez del Constructo para el MMSE y el 3MS: En el MMSE aparecieron tres factores: **Factor I:** Ítems que evalúan nominación, registro, orden de tres comandos, repetición y orientación espacial; **factor II:** Ítems para orientación espacial, evocación, orientación en tiempo y copia; y **factor III:** Conformado por los ítems que evalúan lectura, escritura atención y cálculo matemático.

En el 3MS aparecieron los siguientes factores: **Factor I:** Ítems para escritura, lectura, inversión de operaciones mentales y copia de dos pentágonos. **Factor II:** Ítems que evalúan primera y segunda evocación, orientación temporal, agilidad mental y similitudes. **Factor III:** Ítems para orientación espacial, memoria, lugar y fecha de nacimiento. **Factor IV:** Ítems que evalúan repetición. La tabla 2 muestra la estructura factorial de ambos instrumentos.

Confiabilidad del MMSE y del 3MS:

La correlación de Pearson para las observaciones intra e interobservador en el MMSE fueron: $r = 0.775$ y $r = 0.802$, respectivamente; para el 3MS, la r de Pearson (intraobservador) fue de 0.773 y la r (interobservador) fue de 0.803. Todas ellas estadísticamente significativas ($p < 0.000$)

Confiabilidad de consistencia Interna del MMSE y del 3MS:

El Alfa de Cronbach fue de 0.807, para el MMSE y de 0.882, para el 3MS. Estos índices indican que los instrumentos son confiables.

Al calcular la confiabilidad por concordancia con el diagnóstico clínico, para categorías dicotómicas: sano y deterioro, el coeficiente kappa, fue de 0.392 para MMSE y de 0.537 para 3MS, ambos estadísticamente significativos ($p < 0.000$). Los valores de kappa alcanzados por el MMSE y por el 3MS para tres categorías: sano, deterioro leve y deterioro severo, fueron 0.275 para MMSE y $= 0.505$ para 3MS, ambos estadísticamente significativos ($p < 0.000$).

Tabla 2. Análisis factorial. Matrix rotada de componentes para MMSE y 3MS.

MMSE ÍTEMS	FACTORES			ÍTEMS	3MS			
	I	II	III		I	II	III	IV
Nominación	.876			Escritura	.830			
Registro	.788			Lectura	.763			
Orden tres comandos	.666			Inversión de operaciones mentales	.757			
Repetición	.580			Copia de dos pentágonos	.702			
Orientación espacial	.535	.500		Primera evocación		.869		
Evocación		.826		Segunda evocación		.810		
Orientación en tiempo		.666		Orientación temporal		.589		
Copia		.533		Agilidad mental		.563		
Lectura			.762	Similitudes		.509		
Escritura			.761	modnomin				
Atención y cálculo matemático			.688	Orientación espacial			.801	
				Memoria			.782	
				Lugar y fecha de nacimiento			.678	
				Repetición				.807

Método de extracción: Análisis principal de componentes; método de rotación: Varimax con normalización kaiser. Convergencia obtenida a las 5 iteraciones, en EMA y a las 6 interacciones en 3MS. Los puntos de “quiebre” a partir del cual el número de factores se hace irrelevante son: En MMSE, el factor 3, y en 3MS, el factor 4.

La concordancia entre el diagnóstico clínico y la puntuación del MMSE se empobrece considerablemente al incluir tres categorías diagnósticas. Esto obedece probablemente a la dificultad para tipificar adecuadamente a los sujetos con deterioro cognitivo leve.

Sensibilidad y especificidad:

La sensibilidad y la especificidad del MMSE alcanzaron el valor de 42.0% y 96.3%,

respectivamente.

Sus valores predictivo positivo (VPP) y predictivo negativo (VPN), para MMSE y 3ME, fueron 61.9% y 92.2%, respectivamente, y la exactitud o confiabilidad discriminante 70.3%.

La sensibilidad y la especificidad del 3MS alcanzaron a 68.0% y 85.3%, respectivamente.

Su VPP y su VPN, en ese orden, fueron 39,4% y 95%, y la exactitud o confiabilidad discriminante 77.0%.

El 3MS presentó sensibilidad superior al MMSE, el cual, lógicamente, muestra especificidad mayor. El VPP fue superior para el MMSE y el VPN fue ligeramente superior para el 3MS. La exactitud o confiabilidad discriminante fue ligeramente superior para el 3MS.

Curvas ROC para el MMSE y el 3MS

En la figura 1 se muestran las curvas ROC para el 3MS y el MMSE, se aprecian los valores continuos de los instrumentos en relación con el diagnóstico clínico clasificado en dos categorías: Sin

deterioro y demencia:

Las curvas para ambos instrumentos resultan muy similares, se cruzan a diferentes niveles de las escalas correspondientes e indican que ambos tests poseen una alta sensibilidad. Cuando se someten a contraste las áreas por debajo de las correspondientes curvas (tabla 3), se observa que las dos escalas son relativamente iguales (0.850 vs. 0.881), con errores típicos son bajos y muy similares (0.026 vs. 0.023) y estadísticamente diferentes de 0.500, lo cual indica que las dos escalas son discriminantes. El intervalo de confianza del MMSE está incluido dentro del intervalo de confianza del 3MS, esto permite inferir que las dos áreas no son estadísticamente diferentes entre sí. Es decir, son equivalentes en relación con la validez discriminante.

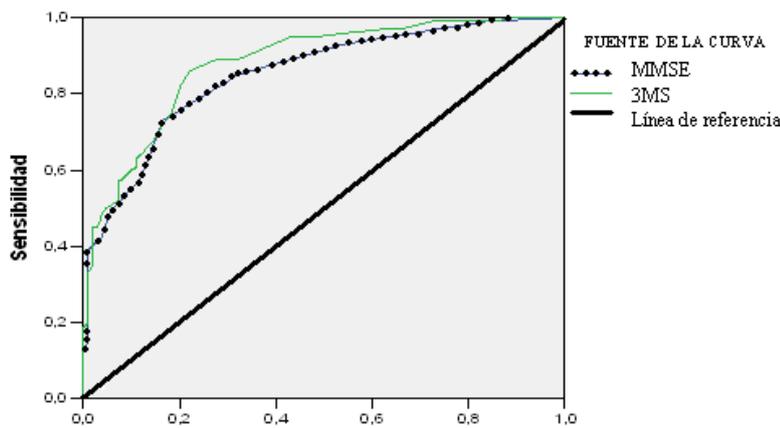


Figura 1. Curvas ROC para el MMSE (en punteado) y el 3MS (línea continua) con relación al diagnóstico clínico en 2 categorías

Tabla 3. Contraste de hipótesis de las áreas por debajo de las curvas ROC para el MMSE y el 3MS

Resultado de la prueba	Área	Error típico (a)	Sign. Asintótica (b)	Intervalo de confianza (asintótico)
MMSE	0.850	0.026	0.000 (***)	0.79 0.90
3MS	0.881	0.023	0.000 (***)	0.83 0.92

(a) = Bajo el supuesto no-paramétrico
 (b) = Hipótesis nula: Área verdadera igual a 0.5
 (***) = Significativo en un nivel alfa inferior a 0.001

Valores de la especificidad y de 1 – especificidad para la escala MMSE, según los diferentes puntos de corte generados por el sistema SPSS (versión 13.0): Entre los puntos de corte 22.5-23.5 la fracción de falsos positivos (1- especificidad) se incrementó considerablemente de 0.009 hasta 0.037, es decir cuatro veces. Con el punto de corte de 23,500 se obtuvo una **sensibilidad** de 0.420 (42.00 %) y una **especificidad** de $1 - 0.037 = 0.963$ (96.30 %). Por lo tanto, cuando la especificidad se incrementa, el punto de inflexión de la curva y el punto de corte propuesto (23) coinciden.

Valores de la especificidad y de 1 – especificidad para la escala 3MS, según los diferentes puntos de corte generados por el sistema SPSS (versión 13.0): Entre los puntos de 60.5-62.0 la fracción de falsos positivos (1- especificidad) se duplica (0.09-0.18). Sin embargo, el punto de corte utilizado en el estudio (79) quedó muy distante de ese punto de inflexión. Este resultado puede explicar el bajo valor predictivo positivo de este instrumento, y

permite justificar la recomendación de que debería disminuirse el punto de corte del mismo.

Análisis inferencial:

Con el propósito de determinar si las variables socio demográficas (género, edad, escolaridad) influyen en forma estadísticamente significativa sobre la el diagnóstico de “deterioro” y “no deterioro”, se calcularon las correspondientes tablas de contingencia y se aplicaron los modelos estadísticos de chi cuadrado y la prueba exacta de Fisher con su correspondiente razón de verosimilitud. En el análisis de los contrastes de hipótesis, las variables edad

y escolaridad resultaron significativas a un nivel alfa del 5 %. El género no resultó significativo. Con relación con la edad, se observó una proporción mayor de casos (deterioro) a los esperados por ley del azar, en los sujetos comprendidos entre 71 y 86 años. En cuanto a la escolaridad, la proporción mayor de casos se presentó en los sujetos de la categoría “hasta primaria incompleta”.

Al evaluar el riesgo relativo indirecto (odds ratio) para el MMSE y el 3MS, en función de las diversas categorías de las variables demográficas, se observó que el sexo tuvo un efecto muy discreto sobre la puntuación en los instrumentos, con una elevación discreta del riesgo de presentar deterioro cognitivo en las mujeres. Este riesgo fue mucho mayor para los sujetos con edad mayor de 71 años y con un nivel de escolaridad de primaria incompleta.

Concordancia entre el MMSE y el 3MS. El índice Kappa para dos categorías: deterioro vs. No deterioro, fue de 0.549 con un error asintótico de 0.057, una aproximación al T de Hotelling de 8.688 con una probabilidad de 0.000 (altamente significativo, a un nivel alfa inferior al 0.001). El valor mínimo aceptable de Kappa es de 0.45; este resultado nos indica que las dos pruebas son concordantes, es decir tienden a medir lo mismo. El número total de sujetos diagnosticados como “deterioro” en el instrumento MMSE fue de 46 (22 %), mientras que en el instrumento 3MS Modificado el número de diagnosticados fue de 84 (40 %).

DISCUSIÓN

El objetivo central del presente estudio fue comparar el MMSE con una versión ampliada del mismo, el 3MS. Se calcularon numerosos indicadores de validez y confiabilidad los cuales se resumen y discuten a continuación.

La validez de contenido fue de 0.8903 (muy buena) para el MMSE y de 0.9571 (excelente) para el 3MS. Este resultado traduce que ambos instrumentos evalúan en forma fidedigna el complejo de síntomas y signos que los expertos denominan “demencia”. Se apreció un índice discretamente superior para el 3MS.

La VCC se evaluó mediante el análisis de correlación entre la puntuación obtenida en cada instrumento y clasificada en tres categorías, y la puntuación obtenida en la escala CDR. Como era de esperarse se obtuvo una correlación negativa, dado que a mayor puntuación en cada instrumento, menor debe ser la ponderación del deterioro cognitivo. El cálculo para el MMSE arrojó un valor del coeficiente de correlación de 0.588 el cual fue ligeramente superior al del 3MS (0.554).

La validez del constructo se evaluó mediante el análisis factorial y reveló un comportamiento diferente para ambos instrumentos. El MMSE arrojó tres factores y el 3MS arrojó cuatro factores. Este resultado sugiere que ambos instrumentos evalúan el nivel de deterioro cognitivo, a través de dimensiones y organizaciones diferentes de sus ítems. Debe recordarse que el 3MS es más extenso que el MMSE, lo cual puede explicar parcialmente la discrepancia en el análisis factorial.

La estabilidad inter observador se realizó mediante el cálculo de la correlación entre la puntuación obtenida por un determinado sujeto al ser evaluado sucesivamente por 2 evaluadores diferentes. La estabilidad intraobservador luego de dos evaluaciones separadas por 48-72 horas por parte del mismo observador (Hernández-Nieto 1995). El coeficiente de correlación arrojó valores superiores a 0.77 en todos los casos, sin diferencias ostensibles entre ambos instrumentos. Esto traduce una alta estabilidad de la evaluación obtenida por ambas pruebas.

La confiabilidad de consistencia interna se evaluó mediante el coeficiente alfa de Cronbach. Este coeficiente para el MMSE, fue de 0.8007 (muy bueno); y para el 3MS, de 0.8820 (muy bueno, tendiendo a excelente). Se pone en evidencia que las diferentes dimensiones de ambos instrumentos están adecuadamente relacionadas entre sí. El 3MS arrojó una confiabilidad discretamente superior al MMSE: La concordancia con el diagnóstico clínico se evaluó mediante el índice kappa. Cuando se utilizaron dos categorías diagnósticas, el índice kappa para el MMSE fue de 0.392, por debajo del el valor mínimo aceptable de 0.45. El 3MS se comportó mejor y obtuvo un índice de 0.537, lo cual se considera “aceptable”.

Cuando se utilizaron tres categorías diagnósticas (sin deterioro, deterioro leve y demencia) el índice kappa disminuyó considerablemente en el MMSE a 0,275 y discretamente en el 3MS a 0.505.

El análisis detallado de los resultados demostró que el porcentaje de sujetos clínicamente definidos como afectados de “deterioro cognitivo leve” y que fueron catalogados como “normales”, de acuerdo a las categorías dicotómicas de los instrumentos, fue de 46.9% con el 3MS y de 85.7% con el MMSE.

La división de los resultados de los instrumentos en 3 categorías (sin deterioro, deterioro leve, demencia) se sustentó en la información publicada para el MMSE (Rosselli et al. 2000). En el caso del 3MS no encontramos estudios publicados que hayan utilizado tres categorías. El punto de corte tradicional (79) distingue demencia de no demencia. En el presente estudio se realizó una categorización arbitraria de la siguiente forma: < 79 = demencia; 80-90 = deterioro leve; > 90 = sin deterioro.

El porcentaje de sujetos clínicamente definidos como portadores de “deterioro cognitivo leve” que fueron clasificados como “sanos” por el MMSE fue de 85.7%. En el caso del 3MS, sólo el 6.1% de sujetos con deterioro cognitivo leve fue catalogado como sano.

En conjunto, estos resultados demuestran que la categoría de deterioro cognitivo leve no es evaluada adecuadamente por ninguno de los instrumentos y aparentemente es sub-diagnosticada por los mismos, particularmente cuando se utilizan sólo dos categorías clínicas (sano vs. demente), como ha sido reportado por Salmon y Butters (1990) y Boeve et al. (2003). Esto explica parcialmente la concordancia baja encontrada entre el diagnóstico clínico y los realizados tanto por el MMSE como por el 3MS, aunque este último se comportó en forma ligeramente superior.

Cuando se utilizan tres categorías el MMSE se mantiene igual pero mejora considerablemente el valor del 3MS (6.1%).

La sensibilidad fue superior para el 3MS (68%) que la del MMSE (42%). Como era de esperarse, por ser valores inversos, la especificidad fue superior para el MMSE (96.3%) que para el 3MS (85.3%). El Valor predictivo positivo, el cual define el porcentaje de sujetos positivos en la prueba y que realmente padecen demencia, fue considerablemente mayor en el MMSE (61.9%) que en el 3MS (39.4%). El Valor predictivo negativo, el cual define el porcentaje de sujetos negativos en la prueba y que realmente no padecen demencia, fue elevado tanto con el MMSE (92.2%) como con el 3MS (95.0%). En conjunto, la exactitud de la

prueba fue ligeramente superior en el 3MS que en el MMSE (77.0% vs. 70.3%).

El bajo Valor predictivo positivo de 3MS está relacionado al comportamiento de su curva ROC. Si bien el área bajo la curva fue ligeramente superior en el 3MS que en el MMSE, la transición brusca de los falsos positivos (1 – especificidad) estuvo cerca del punto de corte usual en el MMSE (Roman et al 1993), pero estuvo muy por debajo del usado para el 3MS. Específicamente, la elevación más brusca de la fracción de falsos positivos ocurrió al pasar el punto de corte de 61 a 62. El punto de corte recomendado por Teng y colaboradores es de 79. En consecuencia, con este nivel de corte se catalogan como positivos muchos sujetos que en realidad son negativos. Esto explica el bajo valor predictivo positivo. Nuestros resultados sugieren entonces que el punto de corte de 23 es adecuado para el MMSE, pero debe disminuirse para el 3MS. Esta sugerencia debe evaluarse en estudios posteriores.

Las variables demográficas influyeron en forma diferencial en los resultados. El sexo no tuvo influencia significativa. Sin embargo, la edad y el nivel de escolaridad sí influyeron considerablemente, tal y como ha sido demostrado por estudios de prevalencia realizados en Europa los cuales mostraron que la demencia aumentaba con la edad (Breteler 1990) y en cuanto a la escolaridad otros estudios mostraron una relación interesante entre bajo nivel educativo y tasas elevadas de prevalencia de la demencia (Fratiglioni et al. s/f, Zhang 1990, Bonaiuto 1990)

REFERENCIAS.

American Psychiatric Association (APA). 1994. DSM-IV Diagnostic and statistical manual of mental disorders Revised. APA, 4ª edición. Washington D.C.

Anthony J, LeResche L, Niaz U et al. 1982 Limits of the mini-mental state as a screening test for dementia and delirium among hospital patients. *Psychology*, 12: 397-408.

Breteler M, Roca W, Amaducci L. 1990. The Prevalence of dementia in Europe: A collaborative study of 1980-1990 findings, *Nueobiol Aging*, 11: 290-291.

Bonaiuto S, Rocca W, Lippi A. et al. 1990. Impact of education and occupation of the prevalence of Alzheimer's disease and multi-infarct dementia in Appignano, Macerata Province, Italy. *Neurology*, 40, Supp 1: 346.

Boeve B, McCormick R, Smith G et al. 2003. Mild Cognitive impairment in the oldest old, *Neurology*, 60: 477-480.

Del Ser T, Morales, J, Bermejo M. 1993. Evaluación del deterioro mental y de la demencia. En: Demencia conceptos actuales: Bermejo F, Del Ser T (Eds.), Madrid. 13-38.

Fischer C. 1951. Senile Dementia: A New Explanation of its Causation, *Can Med. Assoc. J*, 65: 1-7.

Folstein M, Folstein S, McHugh P. 1975. Mini-Mental State Examination: a practical method for grading the cognitive state of patients for the clinician. *Journal Psychiatry Res*, 12: 189-198.

Fratiglioni L, Grut M, Forsell Y et al. s/f. Prevalence of Alzheimer's disease and other dementias in an early urban population. Relationship with sex and education. *Neurology* (en prensa).

Heber R, Brayne C. 1995. Epidemiology of dementia, *Neuroepidemiology*, 14: 240-257

Hernández-Nieto R. A. 1995. El coeficiente de proporción de rangos (Cpr): una alternativa para determinar la validez de contenido y el nivel de concordancia entre jueces en escalas Likert. Trabajo presentado en el XXV Congreso Interamericano de Psicología. San Juan, Puerto Rico: 9 al 14 de Julio.

Hernández Nieto R. 2002. Contribuciones al Análisis Estadístico: El coeficiente de validez de contenido (Cvc) y el Coeficiente Kappa, en la determinación de la validez de contenido según la técnica de juicio de expertos. Producciones Karol, Mérida, Venezuela

Hernández R, Fernández C, Baptista P. 2003. Metodología de la Investigación. McGraw Hill, México.

Jorm A, Korten E, Henderson A. 1987. The Prevalence of dementia: A quantitative integration of the literature. *Act. Psychiat. Scand*, 76: 465-479.

Kukull W, Larson E, Teri L et al. 1994. The Mini-Mental State Examination score and the clinical diagnosis of dementia. *Journal Clin Epidemiol*, 47: 1061-1067.

López Ullibarri G I, Pita Fernández S. 1998. Curvas ROC. *Cad Aten Primaria*, 5: 229-235.

Morris, J. 1993. The Clinical Dementia Rating (CDR). Current version and scoring rules, *Neurology*, 43: 2412-2414.

Rosselli D, Ardila A, Pradilla G et al. 2000. El examen mental abreviado (Mini Mental State Examination) como prueba de selección para el diagnóstico de demencia: estudio poblacional colombiano. *Rev Neurol*, 30: 428-432.

Ruiz A. 2002. Pruebas diagnósticas en medicina clínica: Desarrollo, evaluación y usos. En: Ruiz Morales A, Gómez Restrepo C, Londoño Trujillo D. (Eds.) Investigación Clínica: Epidemiología Clínica

Ramírez et al. 2011. Validez y confiabilidad MMSE y 3MS en diagnóstico demencia. MedULA 20: 128-135

Aplicada. Centro Editorial Javeriano. Bogotá. Colombia.

Salmon D, Thal L, Butters N. 1990. Longitudinal evaluation of dementia of the Alzheimer's Type: a comparison of 3 standardized mental status examinations, *Neurology*, 40: 1225-1230.

Teng, E, Chang H. 1987. The Modified Mini-Mental State (3MS) Examination. *Journal Clin Psychiatry*, 48: 314-318.

Tierney M, Fisher R, Lewis A. 1988. The NINCDS-ADRDA Work group criteria for the clinical diagnosis of probable Alzheimer's disease: A

clinicopathologic study of 57 cases, *Neurology*, 38: 359-364.

Zhang M, Katzman R, Salmon D et al. 1990. The prevalence of dementia of Alzheimer's disease in Shanghai, China: Impact of age, gender and education. *Ann Neurology*, 27: 428-437.

Zauding M. 1992. A New systematic method of measurement and diagnosis of mild cognitive impairment and dementia according to ICD-10 and DSM-III-R criteria, *Int Psychogeriatr*, 4: 203-219.

Recibido: 27 mayo 2011 Aceptado: 16 nov 2011