

Utilidad clínica del Test de Aprendizaje Verbal España Complutense (TAVEC) para detectar Alzheimer y Deterioro Cognitivo Leve en población argentina

Clinical Utility of the Test de Aprendizaje Verbal España Complutense (TAVEC) to discriminate subjects with normal aging, mild cognitive impairment, and Alzheimer disease in Argentinean population

Autora: Marlene Pruvost

Carrera: Lic. en Psicología

Mail: marlene.pru@gmail.com

Resumen

Últimamente se ha puesto gran interés en investigar los sutiles límites que existen en el espectro cognitivo que abarca desde el envejecimiento normal hasta la enfermedad de Alzheimer. El propósito del presente estudio fue determinar la utilidad clínica del Test de Aprendizaje Verbal España Complutense (TAVEC), con el objetivo de identificar marcadores cognitivos de deterioro que sean eficaces, de rápida y económica aplicación, facilitando arribar a diagnósticos tempranos de estas enfermedades neurodegenerativas. Los datos de casos clínicos fueron obtenidos por medio del registro de evaluaciones neuropsicológicas efectuadas entre los años 2012 y 2015 en la Unidad de Neuropsicología y Neurorehabilitación (UNR) del Sanatorio Allende y el Centro Integral de Apoyo Terapéutico y Estimulación Cognitiva (CIATEC). Se conformaron tres grupos muestrales según los diagnósticos clínicos. Utilizando una curva analítica ROC (Receiver Operating Characteristic), se estableció la relación entre los diagnósticos de los pacientes y los puntajes obtenidos por cada uno en el TAVEC. Se seleccionaron los puntos de corte óptimos para identificar a los grupos según sus

puntuaciones. Ante los resultados se puede afirmar que el TAVEC es un test de memoria episódica verbal válido como prueba diagnóstica, con alta capacidad de discriminación entre envejecimiento normal, Deterioro Cognitivo Leve y Demencia.

Abstract

There has been great interest from researchers to investigate the borders between normal aging and Alzheimer disease in the past years. Mild cognitive impairment (MCI) appears as a concept to represent a transitional state between the opposite poles mentioned above, becoming a pathological and prodromal state of dementia. The aim of this study was to establish the clinical utility of the Test de Aprendizaje Verbal España Complutense (TAVEC) to discriminate subjects with mild cognitive impairment, Alzheimer disease and a healthy control group in Argentinean population. The main purpose was to identify efficient, economic and quickly applicable markers of cognitive decline, which will facilitate early diagnosis of neurodegenerative diseases. The cases were recruited from the Sanatorio Allende's Neurorehabilitation Unit database. Three groups were established: 42 subjects with MCI, 44 subjects with dementia and 57 healthy subjects for the control group. The relation between diagnosis and test scores were established using ROC curves, an analytic tool that defines the sensitivity and specificity of a test to discriminate between different populations or groups. The optimal cutting points for detecting dementia and MCI using the TAVEC were selected. The results suggest that TAVEC is valid as a diagnostic test, expressing high discriminative capacity to differentiate between normal aging, MCI and dementia.

Palabras clave: Enfermedad de Alzheimer, Deterioro Cognitivo Leve, marcadores cognitivos, memoria episódica verbal, diagnóstico temprano

Keywords: Alzheimer Disease, Mild Cognitive Impairment, verbal episodic memory, early diagnosis, cognitive markers.

1 - INTRODUCCIÓN

Actualmente las enfermedades neurodegenerativas son un tema de gran impacto en la salud pública, debido al número creciente de adultos afectados y también por los

elevados costos directos e indirectos que implica su tratamiento (Custodio et al., 2012). En Argentina específicamente, la prevalencia de demencia en adultos mayores de 65 años está estimada en 12 a 15 casos por cada 1.000 habitantes. Un estudio piloto (Arizaga et al., 2005) registró deterioro cognitivo en el 23% de los sujetos mayores de 60 años. Según estas cifras se infiere que hay en el país aproximadamente 1.000.000 personas con deterioro cognitivo y 480.000 con Enfermedad de Alzheimer (EA) y otras demencias. Si bien los valores de las tasas de prevalencia e incidencia difieren según la zona geográfica, es significativo que en todas se observa la duplicación en estos valores cada cinco años de edad, oscilando entre 1,5% (en el grupo de 65 a 69 años) y 29.9% (en los de 90 a 94 años) (Centurión, Salera, Anciola, Abriata, Barbieri, Capriati, et al., 2008).

El diagnóstico de la EA es evidente en las fases avanzadas, cuando ya son constatables los trastornos cognitivos y los cambios conductuales. En la práctica clínica, la dificultad radica en realizar el diagnóstico en las fases iniciales o prodrómicas de la EA, teniendo en cuenta que las quejas subjetivas de memoria son muy frecuentes entre las personas mayores que no evolucionan necesariamente a una demencia (Herranz, 2013; Armas Castro, Carrasco García, y Angel Valdés, 2008). Debido a esto es que resulta indispensable disponer de pruebas cognitivas sensibles y específicas que colaboren en la identificación de aquellas personas propensas a desarrollar EA. La mayoría de los autores que han realizado estudios para encontrar predictores neuropsicológicos de progresión hacia EA han enfatizado la necesidad clínica de detectar los signos tempranos de Deterioro Cognitivo Leve (Arnaiz y Almkvist, 2003).

El Deterioro Cognitivo Leve (DCL) ha sido definido como un estadio intermedio dentro del continuum cognitivo que existe entre envejecimiento normal y EA (Petersen, 2004). Se trata de una declinación cognitiva que es superior a lo esperable para la edad de la persona, pero que ocurre sin cambios significativos en el funcionamiento de la vida diaria (Custodio et al., 2012). Aunque no todos los sujetos con DCL evolucionan a EA, el riesgo anual de progresión es mayor para ellos que para la población en general (Custodio et al., 2012).

El diagnóstico de EA se realiza principalmente en base a los criterios del Manual Diagnóstico y Estadístico de Trastornos Mentales cuarta edición (DSM-IV-TR) y del grupo National Institute of Neurological Disorders and Stroke-Alzheimer Disease and Related Disorders (NINCDS-ADRDA). Sin embargo en los últimos años los avances tecnológicos y de investigación han permitido identificar marcadores específicos de EA

lo cual llevó a una revisión de los criterios diagnósticos por parte de un grupo internacional de investigadores expertos en demencia (Dubois et al, 2007).

Esta revisión de criterios permite realizar el diagnóstico en periodos prodrómicos de la enfermedad. El criterio clínico principal es un deterioro temprano de memoria episódica, el cual debe ser acompañado por alguno de los criterios secundarios (biomarcadores) para llegar al diagnóstico de probable de EA. En el 86 a 94% de los casos de EA, el síntoma principal es la amnesia progresiva que se manifiesta como deterioro de la memoria episódica (Westerberg et al., 2013; Dubois et al., 2007). Este deterioro debe ser reportado por el paciente y confirmado por algún informante, pero debe también ser evidenciado objetivamente por medio de pruebas neuropsicológicas estandarizadas.

Tal como explicitan Valls-Pedret, Molinuevo y Romi (2010), existen algunas limitaciones para aplicación clínica de estos nuevos criterios diagnósticos ya que aun no hay consenso sobre cómo interpretar la información que brindan tras su aplicación. No se ha definido con precisión qué resultados de biomarcadores deben ser considerados anormales, así como tampoco se ha explicitado qué pruebas neuropsicológicas se deben aplicar o qué puntos de corte son los más adecuados para definir lo que se considera normal o patológico. Sumado a esto, no todos los centros clínicos disponen de los recursos económicos y logísticos que requiere la aplicación de las técnicas de neuroimagen y análisis de líquido cefalorraquídeo por medio de punción lumbar que posibilitan el estudio de biomarcadores (Valls-Pedret, Molinuevo y Romi, 2010).

En este panorama, se cree que el uso de biomarcadores debe considerarse en casos particulares y bajo el análisis de los expertos para determinar su utilidad. Por su parte, los marcadores cognitivos de deterioro temprano de memoria episódica pueden aplicarse de manera económica y no invasiva mediante la toma de un test neuropsicológico, y ser complementados con la valoración subjetiva de la atrofia temporal medial que evidencie la imagen por resonancia magnética (Alberca Serrano, 2008).

Existen múltiples pruebas que evalúan memoria episódica y que permiten constatar la presencia de déficits en registro, almacenamiento y evocación de información. El patrón típico de compromiso en pacientes con EA se caracteriza por la disminución del recuerdo en sub pruebas de recuerdo libre, que no responde a claves ni a tareas de reconocimiento (Loewenstein, Acevedo, Agron, et al., 2009). Dicho patrón es conocido como perfil hipocampal, el cual está asociado con los principales cambios

patológicos que caracterizan al encéfalo de un paciente con EA: la proliferación de ovillos neurofibrilares y placas seniles en el lóbulo temporal medial, así como también con la disminución del volumen del hipocampo constatada por medio de Resonancia Magnética (RM) cerebral (Dubois et al., 2013, 2007).

Uno de los tests más conocidos de habla hispana para evaluar memoria episódica verbal es el Test de Aprendizaje Verbal España-Complutense (TAVEC), el cual permite la generación de una curva de aprendizaje por medio de múltiples ensayos de recuerdo inmediato (Benedet y Alexandre, 1998). Incluye sub-pruebas de recuerdo libre a corto y a largo plazo, y recuerdo con claves semánticas. Se trata de una prueba ampliamente utilizada en población de habla hispana con demencia así como en otros trastornos neurológicos y psiquiátricos, pero no se han encontrado datos de su validación o adaptación para Argentina.

En el 2013, González-Palau y sus colaboradores han publicado un estudio de validación y utilidad clínica del test de memoria episódica verbal Hopkins Verbal Test-Revised (HVLTR) para detectar EA y DCL en población española. Los responsables de dicha investigación seleccionaron una muestra de 298 sujetos (controles saludables, pacientes con EA y pacientes con DCL) que fueron evaluados con el HVLTR, luego de aplicar la metodología adecuada para traducirlo al idioma español. Se realizó una curva ROC para analizar la sensibilidad y especificidad del test para diferenciar pacientes con DCL y EA a partir de sus puntajes en las pruebas de recuerdo libre y recuerdo demorado. También se identificaron los puntos de corte adecuados para detectar EA en la muestra española, lo cual resulta de gran utilidad para realizar diagnósticos clínicos preliminares a partir de una breve evaluación neuropsicológica. Arribaron a la conclusión de que el HVLTR es un instrumento prometedor para diferenciar entre pacientes con DCL, EA y adultos saludables (Gonzalez-Palau et al., 2013).

En base a lo expuesto anteriormente se destaca la importancia de identificar, mediante el uso de pruebas neuropsicológicas específicas y sensibles a la población regional, diferentes perfiles de deterioro cognitivo amnésico que puedan deberse a enfermedades neurodegenerativas. Se considera que la identificación de aquellas personas con Deterioro Cognitivo Leve es crucial para iniciar acciones diagnósticas y terapéuticas en una fase prodrómica de demencia.

2 - METODOLOGÍA

2.1 – Diseño y Procedimiento

Estudio cuasi experimental de tipo cuantitativo. Ex post facto retrospectivo con grupo cuasi control (León y Montero, 2003). Se realizó un análisis de las puntuaciones obtenidas por adultos mayores de 45 años con quejas mnésicas evaluados por medio del TAVEC que poseían diagnóstico clínico de EA o demencia mixta y también de aquellos que tenían DCL. Asimismo, se recogieron datos de adultos mayores en los que se registró ausencia de deterioro cognitivo. De este modo se establecieron tres grupos de estudio: adultos con demencia, adultos con DCL, y adultos mayores saludables.

2.2 - Participantes

La selección de la muestra se realizó a través de la Unidad de Neuropsicología y Neurorehabilitación (UNR) del Sanatorio Allende y el Centro Integral de Apoyo Terapéutico y Estimulación Cognitiva (CIATEC). Los datos fueron obtenidos por medio del registro de evaluaciones neuropsicológicas efectuadas a adultos que consultaron por fallas mnésicas durante los años 2012 a 2015. Dichas evaluaciones consistieron en una entrevista cualitativa con el paciente y su familia, la aplicación del TAVEC y de cuestionarios conductuales auto-administrados. Asimismo, la evaluación incluyó una batería neuropsicológica completa.

Partiendo de una base de datos inicial de 766 casos, se llegó a la selección de 143 sujetos. Se excluyeron de la muestra todos los individuos con antecedentes de enfermedad cerebro vascular, enfermedad de Parkinson, antecedentes de traumatismos cerebrales, episodio depresivo mayor, anoxia, neurocirugías, infecciones, exposición a tóxicos, sida/VIH y tratamientos con antipsicóticos o antiepilépticos; así como también se excluyeron los casos atípicos. Se conformaron tres grupos muestrales: 42 casos de Deterioro Cognitivo Leve, 44 casos de demencia mixta y EA, y por último un grupo de 57 casos de envejecimiento normal o saludable.

2.3 - Instrumento

El TAVEC (Benedet y Alexandre, 1998) permite la evaluación del aprendizaje verbal y de la memoria episódica. Dispone de una lista inicial (A) de 16 sustantivos,

compuesta por cuatro categorías semánticas distintas. La aplicación consiste en cinco ensayos de aprendizaje inmediato, seguida de un ensayo de aprendizaje de una lista de Interferencia (B). Posteriormente, se presenta al sujeto una segunda prueba de recuerdo libre a corto plazo de la lista A y una prueba de recuerdo libre con claves. Luego de un intervalo de 20 minutos se presenta un ensayo de recuerdo libre demorado, una prueba de recuerdo con claves semánticas y una prueba de reconocimiento.

2.4 - Análisis de datos

Se utilizó el software SPSS 19 para analizar los datos de la muestra seleccionada. En primer lugar se utilizó el análisis ANOVA de un factor (variables numéricas) y Tablas de Contingencia (variables nominales-ordinales) para realizar un análisis descriptivo de los datos demográficos de la muestra. Se realizó un análisis de la asimetría y curtosis de la variable con el objetivo de analizar las características de la distribución muestral. Se realizó la Prueba de Levene para verificar la igualdad de varianzas. Luego se procedió con la aplicación del ANOVA de un factor para determinar si existe un efecto diferencial del diagnóstico clínico sobre las puntuaciones del test de memoria. También se aplicó la prueba d de Cohen para conocer el tamaño de efecto.

Posteriormente, utilizando una curva analítica ROC se estableció la relación entre los diagnósticos de los pacientes y los puntajes obtenidos por cada uno en el TAVEC. A partir de los resultados discriminantes de la curva ROC se seleccionó el puntaje de corte considerado como óptimo para distinguir entre los grupos clínicos en estudio.

3 - RESULTADOS

3.1 - Características socio-demográficas de la muestra

En la Tabla 1 se muestran las características demográficas de la muestra total y también según los grupos diagnósticos establecidos (Saludables, DCL, y Demencia). El 39,9% de la muestra total está conformada por el grupo de sujetos Saludables, mientras que el 29,4% y 30,8% restante corresponden a los grupos de DCL y Demencia

respectivamente. El 63,6% de la muestra total fueron mujeres, predominando en número sobre los hombres en los grupos diagnósticos, con excepción del grupo DCL.

Las diferencias de puntuaciones del Mini Mental (MMSE) entre los grupos dan la pauta del deterioro cognitivo que avanza gradualmente a medida que se va desde el grupo saludable hasta el demencial. Del mismo modo, las diferencias entre las medias de edad de los grupos son significativas. Esta diferencia es particularmente mayor respecto al grupo con demencia, lo cual sugiere pensar que a medida que las personas envejecen tienen mayores probabilidades de padecer deterioro cognitivo o demencia. Entre el grupo control y el grupo con DCL las diferencias de medias de edad son mucho menores (dos años aproximadamente), en comparación con la media de edad en demencia (prácticamente 10 años mayor que en los otros dos grupos). Estos datos son acordes con la documentación científica existente (Peña Casanova, 2005; Arizaga et al., 2005), e indican que a partir de los 65 años de edad hay que prestar más atención a los fallos de memoria para realizar controles y un seguimiento del avance de los mismos.

Tabla 1.
Características demográficas de la muestra completa y según grupos diagnósticos

<i>Características</i>	Total (n=143)	Saludables (n=57)	DCL (n=42)	Demencia (n=44)
Edad (M±SD)	65,88 ± 10,01	61,89 ± 9,47	63,95 ± 8,70	72,89 ± 8,17
MMSE (M±SD)	26,88 ± 4,39	29,25 ± 0,89	27,48 ± 4,53	23,3 ± 4,68
Sexo (%)				
hombre	52 (36,4%)	17 (11,9%)	22 (15,4%)	13 (9,1%)
mujer	91 (63,6%)	40 (28,0%)	20 (14,0%)	31 (21,7%)
Educación (%)				
básica	7 (4,9%)	0 (0,0%)	1 (0,7%)	6 (4,2%)
primario	32 (22,4%)	1 (0,7%)	13 (9,1%)	18 (12,6%)
secundario	44 (30,8%)	26 (18,2%)	10 (7%)	8 (5,6%)
terciario	24 (16,8%)	12 (8,4%)	7 (4,9%)	5 (3,5%)
universitario	36 (25,2%)	18 (12,6%)	11 (7,7%)	7 (4,9%)
TOTAL (%)	100%	39,90%	29,40%	30,80%

3.2 - Análisis de la variabilidad y tamaño de efecto

Se seleccionaron 4 de los 14 puntajes de sub pruebas del TAVEC que se consideran los más indicados para identificar déficits mnésicos de origen neurodegenerativo. Mediante un análisis descriptivo de los datos se observó la asimetría de la muestra, lo cual permitió determinar que la distribución muestral cumple con el

modelo de normalidad mínimo y aceptable para variables de estudio en psicología (± 2 SD). A su vez, tres de las sub pruebas seleccionadas (recuerdo libre a corto plazo, recuerdo libre a largo plazo, y recuerdo inmediato total) obtuvieron una significación superior a 0,05 en la Prueba de Levene, lo cual indica que se cumple la condición de igualdad de varianzas.

Se realizó un ANOVA de un factor para comparar las medias de los grupos y conocer si existe un efecto significativo del diagnóstico neuropsicológico sobre el rendimiento en la prueba de memoria TAVEC (Tabla 2). Las medias de los sujetos en las diferentes sub pruebas varían significativamente según el grupo clínico en el que se encuentren. La significación estadística de $0 < 0,05$ aporta evidencia de que las puntuaciones obtenidas por los pacientes varían significativamente según el diagnóstico clínico que tengan los mismos.

Tabla 2.
Comparación de medias mediante ANOVA de un factor

<i>Sub prueba</i>	<i>Saludable (n=57)</i>		<i>DCL (n=42)</i>		<i>Demencia (n=44)</i>		F	Sig.
	M	SD	M	SD	M	SD		
Recuerdo Inmediato	58,4	(8,48)	37,62	(9,80)	24,67	(9,46)	170,285	0
Recuerdo CP	11,67	(2,51)	5,45	(2,92)	1,45	(2,05)	207,066	0
Recuerdo LP	12,47	(2,50)	5,86	(2,88)	1,43	(2,31)	231,096	0
Discriminación	95,78	(4,10)	85,72	(8,69)	70,39	(13,99)	89,093	0

El contraste Post-Hoc Bonferroni demuestra que el rendimiento mnésico en las sub pruebas del TAVEC es mayor en las personas saludables que en aquellas que tienen DCL y mucho mayor que en aquellas con Demencia. A su vez las personas con DCL tienen un mejor rendimiento mnésico que las personas con Demencia. Utilizando un intervalo de confianza del 99% se encontró que entre todos los grupos la diferencia de medias era significativa al nivel 0.01. Esto podría representarse de la siguiente manera *Saludable > DCL > Demencia*.

Para tener más información al respecto, y conocer con exactitud el tamaño del efecto que tiene el diagnóstico clínico sobre las puntuaciones en la prueba de memoria se realizó la prueba estadística *d* de Cohen. Los resultados se muestran en la Tabla 3. Habiéndose encontrado en todos los casos una *d* > 0.8 se considera que el tamaño del efecto es grande.

Tabla 3.
Medida estandarizada del tamaño de efecto como diferencia de medias

<i>Sub prueba</i>	d de Cohen		<i>DCL- Demencia</i>
	<i>Saludable-DCL</i>	<i>Saludable-Demencia</i>	
Recuerdo Inmediato	2,27	3,76	1,34
Recuerdo CP	2,29	4,48	1,61
Recuerdo LP	2,45	4,6	1,71
Discriminación	1,57	2,8	1,35

3.3 - Especificidad y Sensibilidad diagnóstica de la prueba

En la Figura 1 se observa el gráfico correspondiente a la curva ROC en las sub pruebas del TAVEC para identificar pacientes con Demencia. El área bajo la curva (AUC) en este caso es de 0,997 para la sub prueba de recuerdo a corto plazo (RL-CP), 0,994 para el recuerdo a largo plazo (RL-LP), 0,996 para el recuerdo inmediato (RI-AT) y 0,991 para el índice de discriminación (Discrim). Esto significa que el test posee una excelente capacidad discriminativa para distinguir entre pacientes sanos y enfermos de demencia.

En la Tabla 4 se muestran diferentes puntos de corte en las sub pruebas del TAVEC en base al grado de sensibilidad y especificidad diagnóstica que aportan sobre la muestra clínica. Las puntuaciones iguales o inferiores al punto de corte elegido se consideran indicativas de posible deterioro. Se han distinguido con un sombreado los puntos de corte más próximos a la esquina superior izquierda de la curva, ya que es el punto en que se encuentra el menor número de errores globales: falsos positivos y falsos negativos.

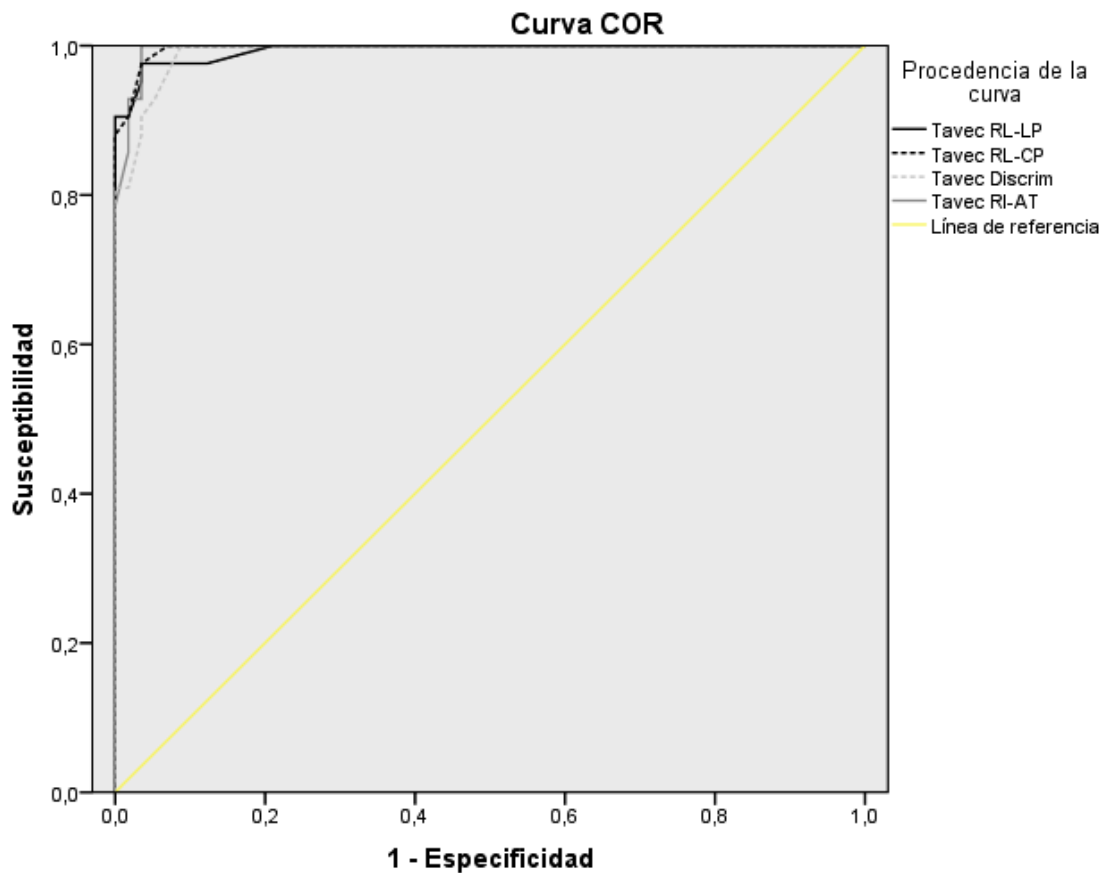


Figura 1. Curva ROC de las puntuaciones del TAVEC para identificar Demencia

Los puntos de corte seleccionados se redondean, ya que las puntuaciones del TAVEC son números enteros (cantidad de palabras recordadas por el sujeto en un ensayo). Como se observa en las Tablas 4 y 5, los puntos de corte arrojados por el sistema SPSS son decimales. Frente a esto se tomó la decisión de redondear ese número decimal según el caso. El único punto de corte que no se redondea es el de Discriminación ya que es un Índice del test que incluye decimales en la puntuación.

Ya que el uso de un punto de corte u otro depende de una decisión clínica en base a la demanda del momento, se puede redondear hacia abajo buscando mayor especificidad, o por el contrario redondear hacia arriba buscando mayor sensibilidad. Se recomienda en caso del DCL redondear hacia arriba como medida preventiva, al igual que en la conversión de DCL a Demencia. En el caso de la identificación de Demencia se puede recomendar redondear hacia abajo buscando mayor especificidad que permita confirmar o complementar el diagnóstico clínico.

Tabla 4.

Sensibilidad y especificidad de las sub pruebas del TAVEC en diferentes puntos de corte para detectar EA y DCL

<i>Subprueba</i>	<i>Deterioro Cognitivo Leve</i>			<i>Demencia</i>		
	<i>Puntaje</i>	<i>Sensibilidad</i>	<i>Especificidad</i>	<i>Puntaje</i>	<i>Sensibilidad</i>	<i>Especificidad</i>
Recuerdo Inmed.	≤46,5	0,881	0,912	≤41,5	0,929	0,965
	≤47,5	0,905	0,912	≤42,5	0,976	0,965
	≤48,5	0,929	0,895	≤43,5	1	0,965
Recuerdo CP	≤7,5	0,81	0,93	≤5,5	0,905	0,982
	≤8,5	0,905	0,912	≤6,5	0,976	0,965
	≤9,5	0,905	0,807	≤7,5	1	0,93
Recuerdo LP	≤8,5	0,857	0,93	≤5,5	0,905	0,982
	≤9,5	0,929	0,877	≤6,5	0,952	0,965
	≤10,5	0,929	0,789	≤7,5	0,976	0,965
Discriminación	≤92,04	0,738	0,825	≤86,33	0,905	0,965
	≤94,31	0,857	0,772	≤87,5	0,929	0,947
	≤95,45	0,905	0,544	≤89,77	1	0,912

El punto de corte óptimo para identificar demencia en las sub pruebas de recuerdo libre a corto plazo es de 6 puntos, el de recuerdo libre a largo plazo es de 7, mientras que para las sub pruebas de recuerdo inmediato y discriminación son de 43 puntos y 87,5% respectivamente. Sin embargo, se incluyen otros puntos de corte en las Tablas ya que la elección de un punto de la curva ROC es una decisión clínica que puede variar teniendo en cuenta cuál es el más adecuado para la población con la que se trabaja o el caso particular del momento.

En la Figura 2 se observa el gráfico correspondiente a la curva ROC en las sub pruebas del TAVEC para identificar pacientes con DCL. En este caso los puntos de corte óptimos son los siguientes: 9 puntos en recuerdo a corto plazo, 10 puntos en recuerdo a largo plazo, 94,31% en discriminación y 48 puntos en recuerdo inmediato. El área bajo la curva (AUC) es de 0,946 para la sub prueba de recuerdo a corto plazo (RL-CP), 0,953 para el recuerdo a largo plazo (RL-LP), 0,945 para el recuerdo inmediato (RI-AT) y 0,873 para el índice de discriminación (Discrim). Esto implica que el test posee una alta capacidad discriminativa para distinguir entre pacientes sanos y enfermos de deterioro cognitivo leve.

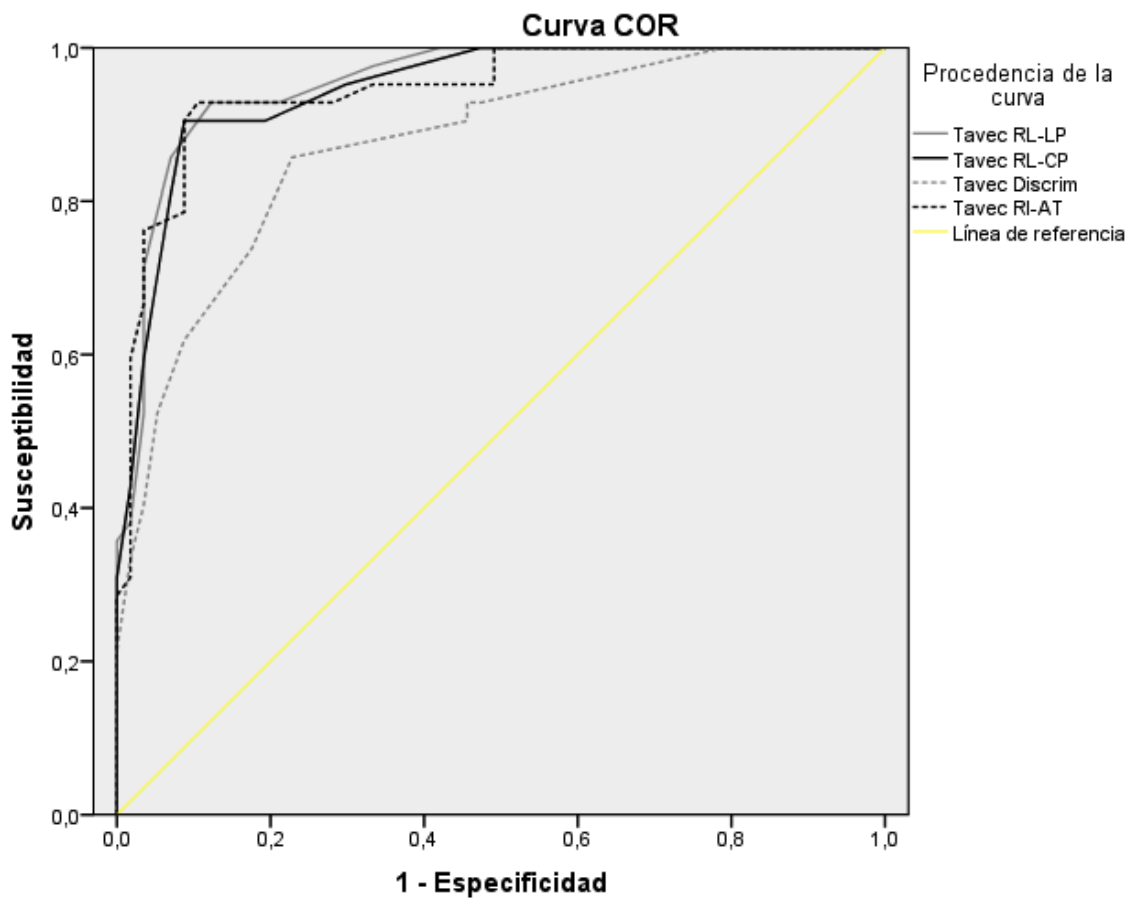


Figura 2. Curva de ROC de las puntuaciones del TAVEC para identificar DCL

La sub prueba de recuerdo a largo plazo es la de mayor utilidad para discriminar entre pacientes con DCL y sujetos sanos, seguida de la sub prueba de recuerdo a corto plazo y la sub prueba de recuerdo inmediato. La sub prueba de discriminación es la que menor AUC posee, siendo la menos discriminativa en comparación a las anteriores.

También se realizó un análisis para identificar sujetos que padecen DCL y avanzan hacia una demencia. En la Figura 3 se observa el gráfico correspondiente a la curva ROC en las sub pruebas del TAVEC para discriminar entre pacientes con DCL y demencia, es decir, para identificar la conversión a demencia. El área bajo la curva (AUC) es de 0,882 para el recuerdo a largo plazo (RL-LP), 0,853 para la sub prueba de recuerdo a corto plazo (RL-CP), 0,846 para el índice de discriminación (Discrim), y 0,83 para el recuerdo inmediato (RI-AT).

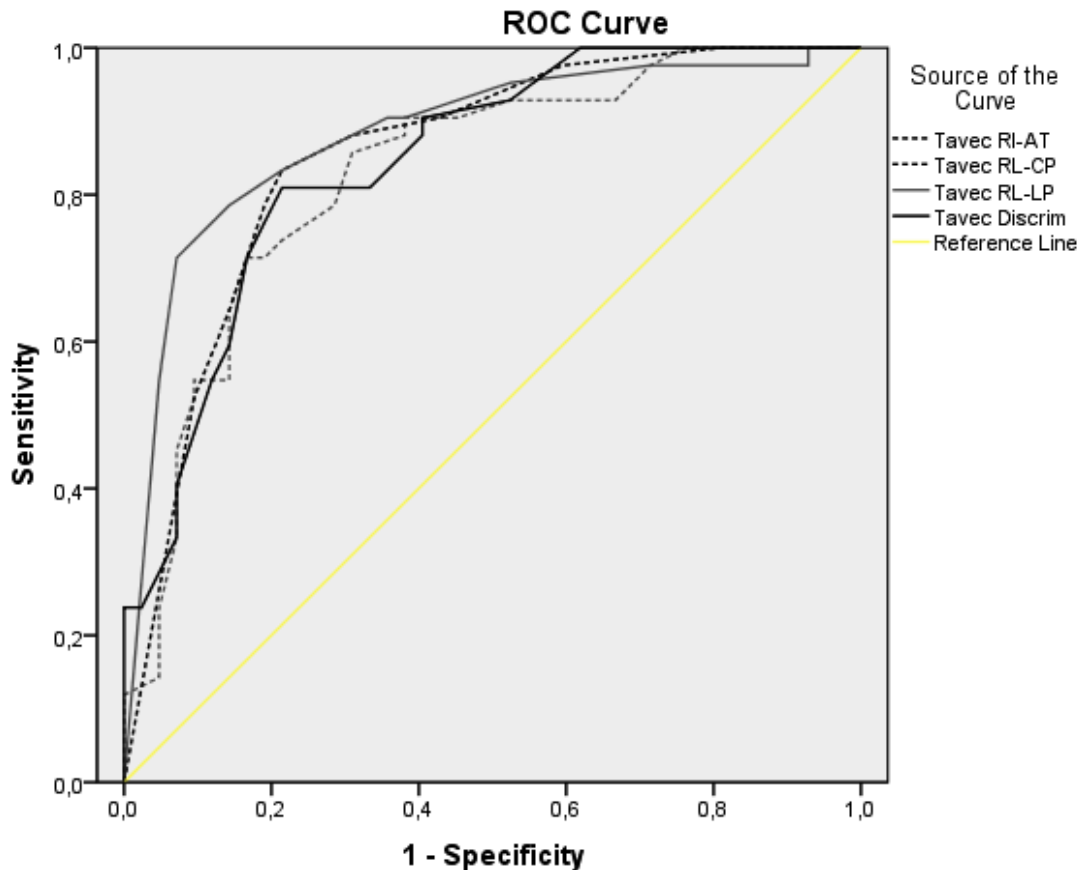


Figura 3. Curva de ROC de las puntuaciones del TAVEC para identificar la conversión a Demencia

Esto implica que para identificar sujetos con DCL que avanzan hacia demencia las pruebas de recuerdo libre a largo plazo y recuerdo libre a corto plazo son las que mayor capacidad de discriminación poseen. Estos resultados sugieren que el TAVEC posee una buena capacidad de discriminación (mayor a 0.8) entre estos grupos clínicos. En este caso los puntos de corte óptimos son los siguientes: 4 puntos en recuerdo a corto plazo, 4 puntos en recuerdo a largo plazo, 80,68% en discriminación y 34 puntos en recuerdo inmediato (Tabla 5).

Se puede observar que la capacidad de discriminación es más elevada para el diagnóstico de Demencia que para el de DCL comparando las AUC en los gráficos de las curvas ROC. A su vez la discriminación para identificar DCL y demencia a partir de la comparación con el grupo saludable (curvas de las Figura 1 y Figura 2) es mucho mayor que la discriminación entre DCL y demencia (curva de la Figura 3). Esto se encuentra directamente relacionado con las dificultades que el continuum cognitivo conlleva para diferenciar entre los cuadros clínicos adyacentes. Entre saludables y

demenciados la brecha es mayor por lo cual la capacidad de discriminación aumenta, pero entre saludables y sujetos con DCL el solapamiento implica mayor dificultad para generar una discriminación clara entre los grupos, por lo cual los valores de sensibilidad y especificidad descienden. Del mismo modo sucede en el solapamiento entre DCL y Demencia, lo cual genera mayor dificultad para detectar la conversión de un estado cognitivo al siguiente. A pesar de esto se lograron altos niveles de discriminación para identificar los grupos diagnósticos utilizando el TAVEC.

Tabla 5.
Sensibilidad y especificidad de las sub pruebas del TAVEC en diferentes puntos de corte para detectar la conversión de DCL en Demencia

<i>Sub prueba</i>	<i>Conversión a Demencia</i>		
	<i>Puntaje</i>	<i>Sensibilidad</i>	<i>Especificidad</i>
Recuerdo Inmed.	≤32,5	0,738	0,786
	≤33,5	0,786	0,714
	≤34,5	0,857	0,69
Recuerdo CP	≤2,5	0,786	0,81
	≤3,5	0,833	0,786
	≤4,5	0,881	0,69
Recuerdo LP	≤2,5	0,786	0,857
	≤3,5	0,833	0,786
	≤4,5	0,905	0,643
Discriminación	≤78,41	0,714	0,833
	≤80,68	0,81	0,786
	≤82,95	0,81	0,667

4 - CONCLUSIONES

El propósito principal de este estudio fue determinar la utilidad clínica del TAVEC para detectar de manera discriminada envejecimiento saludable, Deterioro Cognitivo Leve y Enfermedad de Alzheimer. Se realizó con el objetivo último de identificar marcadores cognitivos de deterioro que sean eficaces, de rápida y económica aplicación, facilitando arribar a diagnósticos tempranos de estas enfermedades neurodegenerativas.

Para lograr el propósito del estudio se analizó la sensibilidad (número de enfermos clasificados como positivos) y especificidad (número de sanos clasificados como negativos) diagnóstica del test de memoria episódica verbal TAVEC mediante el uso de curvas ROC, y se seleccionaron los puntos de corte considerados óptimos para generar diagnósticos clínicos con un margen de error mínimo.

Los valores de sensibilidad y especificidad arrojados por la curva ROC validan al test TAVEC como prueba diagnóstica. Los puntos de corte que se indican como óptimos han sido seleccionados aplicando el criterio más utilizado en pruebas de detección clínica, es decir, aquellos que reflejan mayor equilibrio entre especificidad y sensibilidad. Sin embargo, los puntos de corte seleccionados en este estudio no son unívocos, ya que se pueden tomar otros valores del conjunto seleccionado, los cuales corresponden a diferentes niveles de decisión clínica según la necesidad contextual del momento. Por ejemplo, en una situación en la que se quiera tomar una actitud preventiva (tamizando las posibles enfermedades antes de realizar el diagnóstico) se podría priorizar el valor de sensibilidad sobre el de especificidad. Por el contrario, ante una situación en la que se quiera confirmar un diagnóstico determinado (conjeturado a partir de los resultados en otras pruebas o indicios clínicos) se podría priorizar el valor de especificidad sobre el de sensibilidad.

Para realizar las curvas ROC se seleccionaron 4 de los 14 puntajes que arroja el TAVEC: recuerdo total inmediato (RI-AT), recuerdo libre a corto plazo (RL-CP), recuerdo libre a largo plazo (RL-LP), e índice de discriminación (Discrim). Esta selección se realizó de acuerdo a evidencia que existe en el ámbito de la investigación neuropsicológica. Las medidas de recuerdo demorado (ya sea a corto o a largo plazo) corresponden al dominio más afectado en demencia tipo Alzheimer, y la evidencia sugiere que prácticamente todos los pacientes con esta enfermedad pueden ser discriminados de los controles basándose en los puntajes obtenidos en el recuerdo libre a largo plazo de los test (Zakzanis, Kaplan y Leach, 1999). A su vez, se ha demostrado que los ensayos de recuerdo demorado utilizando el RAVLT contribuyen a la detección temprana de pacientes con EA y a la diferenciación entre EA, DCL y envejecimiento normal (Estévez-González et al., 2003). También se ha encontrado que las puntuaciones de los grupos saludables, con DCL y EA se diferencian en todas las sub pruebas del TAVEC, siendo más prominente en los recuerdos inmediatos y diferidos (Herranz, 2013).

Se realizaron tres curvas ROC utilizando los 4 puntajes descriptos: comparaciones entre el grupo control y DCL, comparaciones entre el grupo control y Demencia, y comparaciones entre DCL y Demencia. Claramente la mayor validez se encuentra en el segundo caso, al comparar saludables y sujetos con demencia la brecha de rendimiento mnésico en el test es mayor por lo cual los índices de discriminación son elevados en las 4 variables. El área bajo la curva (AUC) es superior a 0,99 en todas

ellas, siendo un puntaje casi perfecto ya que el nivel máximo de discriminación diagnóstica se expresa con el número 1.

En el caso de la curva ROC para identificar DCL, aunque los valores de AUC no son tan elevados como para el caso de la demencia, siguen siendo puntajes con una alta capacidad discriminativa. Las variables de recuerdo a corto plazo y recuerdo inmediato poseen un AUC superior a 0,94; mientras que el recuerdo a largo plazo posee el mayor AUC, superior a 0,95. La prueba de discriminación es la que posee menor AUC con un valor de 0,87. Respecto a la tercera curva (que buscaba discriminar el paso de DCL hacia Demencia) se obtuvo una buena capacidad de discriminación con una AUC de 0,882 en la sub prueba de recuerdo libre a largo plazo, seguida por valores de 0,85 a 0,83 en las otras sub pruebas.

En base a los resultados se considera al TAVEC un instrumento de evaluación con excelente capacidad discriminativa al detectar Demencia. Posee a su vez una muy buena capacidad discriminativa para detectar DCL, y una buena capacidad discriminativa para detectar la conversión de DCL a Demencia. Los resultados del estudio son relevantes, ya que aportan conocimiento científico que permitirá utilizar un instrumento de evaluación neuropsicológica para tomar decisiones clínicas de manera económica, rápida y con un margen de error mínimo. Sin embargo, hay algunas consideraciones particulares que deben ser tenidas en cuenta en este estudio. Como toda producción científica, se trata de un trabajo perfectible que puede mejorarse.

Una de las particularidades de este estudio es el hecho de que el grupo muestral de demencia no pudo conformarse de manera pura con casos de Demencia tipo Alzheimer confirmados. Debido a las condiciones epidemiológicas de la población argentina la mayoría de los casos de demencia en el país tienen etiología mixta, es decir neurodegenerativa y vascular. Esto se debe en gran parte a que casi no se realizan procesos prevención primaria ante los factores de riesgo vasculares (sobrepeso, hipertensión arterial, diabetes y tabaquismo como algunos de los principales) que ponen en riesgo al sistema nervioso y contribuyen de manera directa con el deterioro cognitivo de la población a medida que avanza la edad (Doval, 2009; Nigro et al., 1999).

Este hecho implicó la necesidad de incluir en el estudio a algunos casos de demencia con etiología mixta además de los casos de demencia tipo Alzheimer que estaban disponibles, con el objetivo de acceder a un número de casos mínimo aceptable. Todos los casos de demencia incluidos en este estudio poseían signos de neurodegeneración, pero en las resonancias magnéticas de algunos (en los casos de

demencia mixta) también se observaban signos de sufrimiento vascular en el tejido encefálico. A pesar de esta dificultad metodológica en el muestreo, los resultados encontrados parecen no haberse visto afectados, ya que como se describió anteriormente la capacidad de discriminación del test para el grupo de demencia fue casi excelente ($AUC > 0,99$). Sin embargo, es un factor que debería tenerse en cuenta a la hora de utilizar estos resultados en la práctica clínica y en investigaciones futuras.

Otra crítica que se puede realizar a este estudio es que se ha tomado al diagnóstico clínico que realizan los profesionales de neurociencias como referencia para contrastar la validez del instrumento. En el caso de la demencia el diagnóstico clínico suele ser sólido y en este estudio se ha complementado (en todos los casos en que estaba disponible) con los resultados de neuroimagen de los pacientes. Sin embargo, en el caso del Deterioro Cognitivo Leve hay mayores dificultades en cuanto al diagnóstico clínico, ya que a pesar de que en los últimos años su estudio aumenta de manera prominente aún quedan algunos problemas de delimitación que complican la tarea en la práctica. En el presente estudio se utilizaron los criterios diagnósticos de DCL más reconocidos y aceptados globalmente, formulados por Petersen (2004) en la Clínica Mayo. Una posibilidad sería la implementación de múltiples instrumentos o estudios que se complementen para darle mayor solidez al diagnóstico clínico de DCL.

Como líneas de investigación futura para profundizar en la temática se considera de gran importancia la realización de baremos para población argentina del TAVEC, ya que actualmente no se cuenta con los mismos y se trabaja con los originales españoles. La normalización y validación completa de instrumentos de evaluación neuropsicológica a la población argentina es sin duda una necesidad si se pretende seguir avanzando en esta área de estudio. También se considera que la implementación de estudios longitudinales puede ayudar a clarificar los solapamientos entre las entidades clínicas del continuum cognitivo, principalmente entre DCL y las otras dos condiciones estudiadas. El seguimiento a largo plazo del rendimiento cognitivo de los participantes con pruebas como el TAVEC puede brindar información clave. Este tipo de investigaciones ya se han realizado con resultados interesantes en otros países como España, pero ciertamente implican de un presupuesto considerable, difícilmente realizable sin apoyo de instituciones estatales o privadas. Por último, se considera que la realización de estudios que complementen el uso de herramientas de evaluación neuropsicológica con biomarcadores sería de gran utilidad para avanzar hacia un futuro

en que los diagnósticos tempranos de las enfermedades neurodegenerativas sean predominantes.

En conclusión, los resultados encontrados en este estudio son coherentes con la literatura científica precedente y cumplieron con las expectativas de lo que se esperaba lograr con este proyecto. Ante estos resultados se puede afirmar que el TAVEC es un test de memoria episódica verbal válido como prueba diagnóstica, con una alta capacidad de discriminación entre envejecimiento normal, Deterioro Cognitivo Leve y Demencia, siendo su mayor potencial discriminatorio entre controles y pacientes con demencia. Se trata de un instrumento de gran utilidad clínica, que además posee la ventaja de que su aplicación es rápida y económica.

A partir de esto se espera que se tengan en cuenta algunos aspectos a la hora de utilizar estos resultados. En primer lugar, se sugiere que el TAVEC sea administrado por personas capacitadas adecuadamente para evaluar del rendimiento mnésico, con un entrenamiento clínico que le permita reconocer variables que pueden intervenir en los resultados de una prueba: fatiga, déficit atencional, desinterés, etc. En segundo lugar, se considera de gran importancia aclarar que encontrar un único puntaje inferior al punto de corte seleccionado en alguna de las sub pruebas no es suficiente para determinar si la persona padece o no de deterioro cognitivo. Se recomienda que se evalúen los 4 puntajes seleccionados en esta investigación contemplando que al menos dos de ellos deben encontrarse por debajo de lo esperable para tener mayor seguridad que no se trate de un resultado azaroso. También deberá contemplarse la edad que tenga el sujeto y aspectos cualitativos, como por ejemplo la curva de aprendizaje verbal que haya tenido el evaluado en los primeros cinco ensayos de recuerdo inmediato.

Así mismo, incluso habiendo encontrado un rendimiento globalmente negativo en la prueba, ese resultado no puede considerarse de manera aislada como indicador suficiente para determinar un diagnóstico. Se trata, tal como fue planteado en los objetivos del estudio, de una herramienta que busca facilitar la detección de casos tempranos para alertar sobre el posible avance hacia una enfermedad neurodegenerativa. En todo caso se deberá tomar este indicio para continuar con una serie de estudios exhaustivos que permitan o no confirmar el diagnóstico preliminar.

REFERENCIAS

- Alberca Serrano, R., (2008). ¿Qué marcadores clínicos y paraclínicos predicen la conversión del deterioro cognitivo leve en una demencia? En D. A., Pérez Martínez y A., BerbelGarcía (eds.), *60 Cuestiones Prácticas en Neurogeriatría* (pp. 15-17). Madrid: SANED.
- Arizaga, R.L., Gogorza, R.E., Allegri, R.F., Barman, D., Morales, M.C., Harris, P. y Pallo, V., (2005). Deterioro cognitivo en mayores de 60 años en Cañuelas (Argentina). Resultados del piloto del Estudio Ceibo *Rev.Neurol.Arg.*; 30 (2): 83-9
- Armas Castro, J., Carrasco García, M. R., y Angel Valdés, S. (2008). Los factores de progresión del deterioro cognitivo leve a la enfermedad de Alzheimer. *Revista Habanera de Ciencias Médicas*, 7(4), 0-0.
- Arnáiz, E., y Almkvist, O. (2003). Neuropsychological features of mild cognitive impairment and preclinical Alzheimer's disease. *Acta Neurologica Scandinavica*, 107(s179), 34-41.
- Benedet, M. J., y Alejandre, M. A. (Eds.). (1998). *TAVEC. Test de Aprendizaje Verbal España-Complutense*. Madrid: TEA.
- Custodio, N., Herrera, E., Lira, D., Montesinos, R., Linares, J., y Bendezú, L. (2012). Deterioro cognitivo leve:¿ dónde termina el envejecimiento normal y empieza la demencia?. In *Anales de la Facultad de Medicina* (Vol. 73, No. 4, pp. 321-330). UNMSM. Facultad de Medicina.
- Doval, H. C. (2009). Más allá de la prevención cardiovascular. En el camino hacia la promoción de la salud. *Revista Argentina de Cardiología*, 77(4), 335-341.
- Dubois, B., Epelbaum, S., Santos, A., Di Stefano, F., Julian, A., Michon, A., et al., (2013). Alzheimer disease: from biomarkers to diagnosis. *Rev Neurol (Paris)*, 169(10), 744-751.
- Dubois, B., Feldman, H. H., Jacova, C., Dekosky, S. T., Barberger-Gateau, P., Cummings, J., et al., (2007). Research criteria for the diagnosis of Alzheimer's disease: revising the NINCDS-ADRDA criteria. *LancetNeurol*, 6(8), 734-746.
- Estévez-González, A., Kulisevsky, J., Boltos, A., Otermín, P., y García-Sánchez, C. (2003). Rey verbal learning test is a useful tool for differential diagnosis in the preclinical phase of Alzheimer's disease: comparison with mild cognitive impairment and normal aging. *International journal of geriatric psychiatry*, 18(11), 1021-1028.

- Gonzalez-Palau, F. et al., (2013). Clinical utility of the Hopkins Verbal Test-Revised for detecting Alzheimer's Disease and Mild Cognitive Impairment in Spanish population. *Archives of Clinical Neuropsychology*. Downloaded from <http://acn.oxfordjournals.org/>
- Herranz, S. G. (2013). Relevancia de los test neuropsicológicos de evaluación de la memoria episódica y de la función ejecutiva como marcadores en la detección precoz del Deterioro Cognitivo Ligero (DCL). Seguimiento longitudinal (Doctoral dissertation, Universidad Nacional de Educación a Distancia).
- Centurión, E., Salera, C., Anciola, J., Abriata, M. G., Barbieri, M. E., Capriati, A. J., et al. (2008). Estado de conocimiento y agenda de prioridades para la toma de decisiones en Enfermedades Neurológicas: Alzheimer y otras Demencias en Argentina. Buenos Aires: Academia Nacional de Medicina
- León, O. G., y Montero, I. (2003). *Métodos de investigación*. Madrid: McGraw Hill.
- Loewenstein, D. A., Acevedo, A., Agron, J., Ownby, R., Barker, W., Strauman, S., et al. (2009). The utility of a brief memory screen in the diagnosis of mild memory impairment in the elderly: preliminary study. *Am J Geriatr Psychiatry*, 17(5), 437-440.
- Nigro, D. A. N. T. E., Vergottini, J. C., Kuschnir, E. M. I. L. I. O., Bendersky, M. A. R. I. O., Campo, I. V. O. N., De Roiter, H. G., & Kevorcof, G. R. E. G. O. R. I. O. (1999). Epidemiología de la hipertensión arterial en la ciudad de Córdoba, Argentina. *Rev Fed Arg Cardiol*, 28, 69-75.
- Petersen, R. C. (2004). Mild cognitive impairment as a diagnostic entity. *Journal of internal medicine*, 256(3), 183-194.
- Peña Casanova, J. (2005). *Activemos la mente, actualización 2005*. Barcelona, Fundación la Caixa.
- Valls-Pedret, C., Molinuevo, J. L., y Rami, L. (2010). Diagnóstico precoz de la enfermedad de Alzheimer: fase prodrómica y preclínica. *Rev Neurol*, 51(8), 471-480.
- Westerberg, C., Mayes, A., Florczak, S. M., Chen, Y., Creery, J., Parrish, T., ... & Paller, K. A. (2013). Distinct medial temporal contributions to different forms of recognition in amnesic mild cognitive impairment and Alzheimer's disease. *Neuropsychologia*, 51(12), 2450-2461.

Zakzanis, K., Kaplan, E. F., Leach, L. (1999). *Neuropsychological Differential Diagnosis*. Sweat & Zeitlinger.