

TESIS DOCTORAL

DEMENCIA ASOCIADA A ENFERMEDAD DE ALZHEIMER Y OTRAS

DEMENCIAS: EVIDENCIAS DIAGNÓSTICAS

INGRID ARÉVALO RODRÍGUEZ

DIRECTORES: XAVIER BONFILL COSP

PABLO ALONSO-COELLO

BARCELONA, MAYO 2015

UNIVERSITAT AUTÒNOMA DE BARCELONA

DEPARTAMENT DE PEDIATRIA, d'OBSTETRICIA I GINECOLOGIA, I DE MEDICINA PREVENTIVA

TESIS DOCTORAL

DEMENCIA ASOCIADA A ENFERMEDAD DE ALZHEIMER Y OTRAS DEMENCIAS: EVIDENCIAS
DIAGNÓSTICAS

DOCTORANDA

INGRID AREVALO RODRIGUEZ

DIRECTOR

XAVIER BONFILL COSP

DIRECTOR

PABLO ALONSO COELLO

BARCELONA, MAYO 2015

INGRID AREVALO RODRIGUEZ

Psicóloga. Epidemióloga Clínica.

Instructor Asociado División de Investigaciones. Facultad de Medicina.

Fundación Universitaria de Ciencias de la Salud

Bogotá, Colombia

TABLA DE CONTENIDO

1. RESUMEN	5
2. ABSTRACT	7
3. INTRODUCCION	9
4. OBJETIVOS	24
5. METODOLOGIA.....	26
4. RESULTADOS	41
5. DISCUSION.....	86
6. CONCLUSIONES	93
7. REFERENCIAS.....	94

1. RESUMEN

Antecedentes: La demencia es una enfermedad causada por un importante número de patologías, las cuales afectan diferentes dominios cognitivos e interfieren en las actividades de la vida diaria. El uso de herramientas diagnósticas en demencia se ha perfeccionado en los últimos años para poder identificar los pacientes con alto riesgo de desarrollar la enfermedad, y así ofrecer un tratamiento que detenga su progresión o corregir las causas que la originan. Para las diferentes herramientas diagnósticas, incluso aquellas que se han vuelto tradicionales en la práctica médica, es indispensable que se conozca su verdadero valor en el diagnóstico y progresión de esta enfermedad. La evaluación sistemática de su capacidad diagnóstica debe responder a los retos clínicos de la actualidad en el abordaje de estos pacientes.

Objetivo: Determinar cuál es la calidad de la evidencia referente a las herramientas diagnósticas empleadas para el abordaje de los pacientes con demencia por Enfermedad de Alzheimer y otras demencias.

Métodos: Se desarrollaron tres trabajos de investigación. En el primero se evaluó la calidad de las guías de práctica clínica que abordaron el diagnóstico de la demencia. En el segundo se evaluó la calidad de las revisiones sistemáticas de características operativas de instrumentos para el diagnóstico de demencia por EA y otras demencias. En el tercer trabajo se determinó la exactitud diagnóstica del MMSE para la detección de sujetos con DCL quienes pueden progresar a demencia.

Resultados: En cuanto a la calidad de las guías de práctica clínica en el diagnóstico de demencia, encontramos que 40% de las GPC evaluadas fueron consideradas de alta calidad; entre ellas, se hizo uso al menos de cuatro diferentes sistemas para presentar las recomendaciones diagnósticas, siendo heterogénea la presentación de la calidad de la evidencia subyacente y la fuerza de la recomendación. En relación a la calidad de las revisiones sistemáticas, se identificaron 24 estudios, incluyendo solo 10 de ellos valoración de la calidad de los estudios primarios, presentando también deficiencias destacadas de reporte y metodológicas. Por último, encontramos 11 estudios altamente heterogéneos y riesgo de sesgo considerable, lo cual impidió la síntesis numérica de la información.

Conclusiones: 1) La calidad de las GPCs que abordan el diagnóstico de demencia por EA y otras demencias es heterogénea en su calidad metodológica, y las recomendaciones dirigidas al diagnóstico varían ampliamente en términos de la valoración de la calidad de la evidencia subyacente y la confianza en la misma. 2) Las revisiones sistemáticas de las herramientas diagnósticas en demencia publicadas hasta 2013 presentan deficiencias metodológicas importantes, que impiden en su mayoría la diseminación de sus conclusiones en la práctica clínica. 3) No existe evidencia que apoye el uso extendido del MMSE para determinar la posible conversión de los pacientes con DCL a demencia.

2. ABSTRACT

Background: Dementia is a disease caused by an important number of pathologies, which affect several cognitive domains and interfere with daily activities. Use of diagnostic tools in dementia has change in recent years, in order to identify patients at high risk and then provide adequate management. For all diagnostic tests, including those that have become traditional in clinical practice, it is important to know its role and value in diagnosis and progression of dementia. Systematic assessment of diagnostic accuracy for tests in dementia should be helpful in this task.

Objective: To determine the quality of the evidence regarding the diagnostic tools used for identification and follow-up of patients with Alzheimer's disease dementia and other dementias.

Methods: Three research projects were developed. In the first project, we appraised the quality of clinical practice guidelines (CPG) focused on diagnosis of dementia. In the second project we assessed the quality of systematic reviews in dementia focused on diagnostic test accuracy. Finally, in the third project we assessed the accuracy of MMSE for detection of subjects with MCI who may progress to dementia.

Results: Regarding the quality of CPG focused on diagnosis of dementia, we found that only 40% of them were considered of high quality. For those guidelines, at least four different systems to developed recommendations were used. Regarding the quality of diagnostic systematic reviews, we identified 24 studies; only ten studies assessed the quality of primary studies appraised and we identified several shortcomings in reporting and in its methodological quality. Finally, we found eleven heterogeneous studies related to MMSE to assess conversion from MCI to dementia, but a metanalysis were not performed due to the variability of data.

Conclusions: 1) Quality of GPC addressing diagnosis of dementia is heterogeneous and their recommendations vary in terms of the confidence in evidence and the impact in management of people with dementia. 2) Systematic reviews of diagnostic tests in dementia have significant methodological shortcomings which preclude their dissemination in clinical practice. 3) At present

there is no evidence to support the use of MMSE to determine the conversion of people with MCI to dementia.

3. INTRODUCCION

La demencia es una enfermedad causada por un importante número de patologías, las cuales afectan diferentes dominios cognitivos como la memoria, el lenguaje, el comportamiento y las funciones ejecutivas, e interfieren en las actividades de la vida diaria (1). La demencia de origen neurodegenerativo se caracteriza por pérdida neuronal y sináptica y por el depósito cerebral de agregados proteicos insolubles intra y/o extracelulares, dentro de las cuales merece citarse la demencia por Enfermedad de Alzheimer (demencia por EA), la demencia por cuerpos de Lewy (DL), la demencia asociada a Enfermedad de Parkinson (demencia por EP) y la demencia frontotemporal (DFT) (2). En otras oportunidades, el compromiso es de origen vascular, produciendo la demencia vascular (DV), la cual ocupa el segundo lugar después de la demencia por EA. Otras demencias incluyen la infección por VIH o por partículas proteicas infecciosas denominadas priones, como la demencia de Creutzfeldt-Jakob de tipo esporádico o la nueva variante denominada enfermedad de la vaca loca (CJnv). También existen otras causas como las producidas por efectos tóxicos, de origen endocrino, metabólicas, secundarias a vasculitis o a traumatismos.

Si bien se espera que con la edad se produzcan cambios cognitivos, la demencia no forma parte del denominado “envejecimiento normal” y la mayoría de los adultos mayores no experimentan esta condición en ningún momento de su vida (1). Sin embargo, las más recientes cifras de esta condición a nivel mundial, apuntan a que el número de personas con demencia ha venido aumentando, y se espera un incremento de la prevalencia para el 2030 al 15% (hasta 75.62 millones de habitantes) y al 17% para el 2050 (hasta 135.46 millones de afectados)(3, 4), con un impacto en el gasto en salud superior a 422 billones de dólares (5, 6). Este panorama hace de la demencia una prioridad de la salud pública a nivel mundial teniendo en cuenta los niveles crecientes de envejecimiento de la población, y representa diferentes retos para determinar su adecuada prevención, diagnóstico y manejo.

En el año 2001 se introdujo el término deterioro cognitivo leve (Mild cognitive impairment, DCL), el cual hace referencia a un estadio prodrómico previo a la demencia, el cual representa una oportunidad de intervención y manejo temprano de los pacientes para evitar su eventual

progresión a la demencia. Se ha estimado que la prevalencia del DCL varía entre 10 a 20% en adultos mayores de 65 años, mientras que en personas mayores de 70 años oscila entre 14 a 18%, siendo estas cifras altamente dependientes de los criterios diagnósticos empleados para su estimación (7, 8).

A continuación se abordarán los principales aspectos relacionados con la demencia, en especial la demencia por EA, su diagnóstico, curso clínico y clasificación, siendo estos aspectos fundamentales para el manejo y seguimiento.

TIPOS DE DEMENCIA

Como se mencionó previamente, la demencia de origen neurodegenerativo se refiere principalmente a la demencia **por enfermedad de Alzheimer (EA)**, la cual se define como el deterioro cognitivo en más de un dominio siendo uno de ellos la alteración de la memoria de tipo hipocampal, e interferencia en las actividades de la vida diaria (9-11). En la mayoría de casos la presentación de la demencia por EA es esporádica, aunque existen casos familiares en menos del 2% asociados a mutaciones conocidas con transmisión autosómica dominante (2).

Por otra parte, la **demencia vascular (DV)** es la segunda causa de demencia en los ancianos y representa un rango de condiciones de compromiso cognitivo asociados con diferentes factores de riesgo vascular como HTA, diabetes, dislipidemia, sobrepeso y tabaquismo, los cuales pueden llevar al paciente a infartos cerebrales en la gran mayoría de los casos (80%) o a hemorragia cerebral (20%) (12). Puede comprometer grandes arterias o producir enfermedad isquémica de pequeño vaso (13). Una de las dificultades en el diagnóstico de esta demencia consiste en la diferenciación de la demencia vascular con la demencia por EA ya que la sola aproximación clínica puede no bastar para la adecuada clasificación, e incluso se conoce que ambas formas de demencia pueden coexistir en determinados casos conformando una demencia de tipo mixto (14).

La demencia por EA, la de tipo vascular y más recientemente la asociada a EP, puede presentar un estadio prodrómico previo denominado **deterioro cognitivo leve**, caracterizado por un deterioro en la función cognitiva que no interfiere con las actividades de la vida diaria (8, 15, 16). En dichos

estadios, los pacientes presentan alteración cognitiva con queja subjetiva de memoria anormal para la edad, pero aún no tienen demencia porque conservan normal la función de las actividades esenciales. Si el compromiso de la memoria se presenta aisladamente, se denomina DCL de tipo amnésico, pero también puede ser DCL no amnésico con compromiso de otro dominio o DCL multidominio. Algunos pacientes con DCL progresan a la demencia mientras que otros pueden permanecer con DCL por varios años o regresar a una cognición normal para su rango de edad (17).

La demencia por **cuerpos de Lewy (DL)** produce deterioro cognitivo de naturaleza fluctuante, con alucinaciones visuales y síntomas de parkinsonismo (18, 19). Es la tercera causa más frecuente de demencia después de la Enfermedad de Alzheimer y las causas vasculares. Esta alteración tiene componentes corticales y subcorticales que se evidencian en la valoración neuropsicológica de los pacientes con mayor compromiso de la atención, funciones viso-espaciales y disfunción ejecutiva. Se conoce que el déficit de procesamiento visual es el causante de las percepciones deformadas propias de esta condición (19). También son frecuentes los signos de disfunción autonómica, el trastorno de la conducta del sueño REM, la hipersomnia, las caídas y la hipersensibilidad a los neurolépticos (2). Una forma cercana a la demencia por CL es la **demencia asociada a la Enfermedad de Parkinson (EP)**, la cual se presenta de manera más tardía, en general un año después de los síntomas motores, afectando varios dominios como la atención, la función ejecutiva y la habilidad viso-espacial (20). La demencia asociada a EP posee grandes similitudes con la DLB, incluso en su fisiopatología, presentándose lesiones anatomopatológicas similares para ambos casos en los estudios de autopsias (21).

Finalmente, la **demencia frontotemporal (DFT)** es un proceso neurodegenerativo focal que produce atrofia selectiva de los lóbulos temporales y frontales (22). Si bien la demencia frontotemporal es menos frecuente que las otras demencias previamente expuestas, es la segunda demencia neurodegenerativa más común en personas menores de 65 años después de la Enfermedad de Alzheimer. Se han identificado varios tipos, entre ellos la variante comportamental (DFTvc) y las variantes que comprometen el lenguaje: afasia primaria progresiva (APP). Esta última se identifica específicamente como variante afasia no fluente/agramática, demencia semántica (DS) y logopénea (2). En general los pacientes presentan cambios comportamentales con cambios

conductuales, alteración del pensamiento o deterioro en el lenguaje (producción o comprensión), que pueden ser interpretados como desórdenes de origen psiquiátrico (2, 22).

ETIOLOGIA Y PATOGENESIS DE LA DEMENCIA

Las causas que desencadenan la aparición de la demencia por EA se encuentran en investigación debido a que han surgido muchas hipótesis respecto al papel que juega la proteína β -amiloide o la proteína Tau. Los mecanismos que producen el daño cerebral y su expresión sintomática final aun no han sido completamente explicados (1). En las formas familiares de la EA se han identificado mutaciones genéticas en APP, PS1 y PS2, así como la presencia del gen de APOE₄ en el cromosoma 19 y el gen MTHFD1L en el cromosoma 6 en las formas esporádicas, llevando a la conclusión de que la enfermedad puede tener diferentes orígenes (17, 23). Si bien la causa de la demencia por EA es desconocida, los rasgos histopatológicos de la enfermedad incluyen la presencia de placas neuríticas, sumadas a evidencia de niveles alterados de Proteína Tau, así como ovillos neurofibrilares (24). Durante más de 20 años se ha postulado que los eventos patogénicos iniciales de esta demencia involucran la presencia de péptido β - Amiloide depositado en la parénquima cerebral, y que la ausencia de balance entre la producción y procesamiento produce como consecuencia final la demencia por EA (25). En 2013, Jack y colaboradores actualizaron su modelo fisiopatológico de la demencia por EA, el cual involucraba la participación de diferentes biomarcadores asociados con el depósito de β - Amiloide y asociados a neurodegeneración, los cuales a su vez podrían explicar la aparición clínica de la enfermedad (26). El modelo propone una secuencia de los biomarcadores con tendencia al declive en relación al tiempo, con desenlaces dependientes de la edad del paciente, los factores de riesgo asociados y el tiempo de evolución de la condición (26). A diferencia del modelo expuesto en el 2010, Jack y colaboradores han subrayado en su actualización que las alteraciones en el B- Amiloide y aquellas relacionadas con la Proteína Tau son procesos fisiopatológicamente independientes, siendo ambos posibles causantes de la demencia. Si bien el modelo cuenta con evidencia indirecta para sustentar su validez, se requieren estudios de cohortes con amplio seguimiento que permitan evaluar la evolución real de los biomarcadores propuestos (26).

Por otra parte, la fisiopatología de la demencia vascular se asocia con la presencia de lesiones cerebrovasculares, siendo su extensión y severidad correlacionadas con la magnitud del daño cognitivo subsecuente (14). Los infartos cerebrales representan regiones del cerebro que han tenido disminución de perfusión cerebral en territorio de pequeño o gran vaso las cuales pueden ser observadas en RM cerebral simple o, en algunos casos, en el examen histopatológico (14, 27). El número o volumen de pérdida necesario para generar deterioro cognitivo no ha sido aún establecido; algunas regiones como el Tálamo, los núcleos basales y el Giro son más vulnerables y corresponden a áreas estratégicas del cerebro para desarrollar demencia vascular (14). Sin embargo, varios estudios sugieren que el síndrome multi-infarto lleva con mayor facilidad a desarrollar demencia vascular, teniendo en cuenta que se producen mayores cambios fisiopatológicos relacionados con hipoxia difusa, inflamación, estrés oxidativo o ruptura de la barrera hematoencefálica (14).

La fisiopatología de la demencia por cuerpos de Lewy incluye la degeneración del citoplasma neuronal derivado de la acumulación anormal de proteínas (alfa-sinucleína, además de ubiquitina y proteínas de neurofilamento) (28). La presencia de placas neurofibrilares o daño neuronal puede ocurrir, pero no es exclusivo de esta patología. Las pruebas de inmunohistoquímica son las más adecuadas para la correcta clasificación de esta demencia, pero estas sólo pueden hacerse posterior a la muerte del paciente (19, 28). Para muchos autores, la demencia asociada a EP comparte similitudes fisiopatológicas con DL, siendo motivo de debate cuales podrían ser las diferencias fundamentales entre los dos síndromes (29).

La demencia frontotemporal, por su parte, representa la demencia más heterogénea en términos de etiología, genética y fisiopatología (22). La presencia de inclusiones neuronales y gliales determina el tipo de esta demencia: tau positiva (isoformas 3R), ubiquitina tau-negativa o sin inclusiones. La observación macroscópica revela atrofia cortical de los lóbulos frontales y temporales dilatación ventricular y atrofia del núcleo caudado. En la afasia primaria progresiva y en la demencia semántica, la atrofia se localiza en la región perisilviana izquierda y en la amígdala, en la corteza orbitofrontal, temporal anterior e insular anterior; En la DFT con variante comportamental la atrofia se localiza en la región dorsolateral- frontal derecha, con compromiso del cíngulo y de la corteza insular anterior (30).

CRITERIOS DIAGNÓSTICOS PARA LA DEMENCIA

En la actualidad no existe un patrón de oro desde el punto de vista clínico, neuroimágenes o de biomarcadores para establecer el diagnóstico de los diferentes tipos de demencia y solamente la evidencia histopatológica puede confirmarla en algunos casos. Por este motivo también los hallazgos neuropatológicos han sido cuestionados (31). Existen criterios clínicos para el diagnóstico de la demencia, los cuales han sido diseñados por diferentes grupos de trabajo y publicados en documentos como el Manual Diagnóstico y Estadístico de Enfermedades Mentales (DSM-IV y V publicado en el 2014 y que ya viene siendo utilizado por algunos grupos)(10, 32), así como aquellos presentados por la Clasificación Internacional de Enfermedades (ICD-10) (33), entre otros. En la tabla 1 se presentan dichos criterios.

Tabla 1. Criterios diagnósticos generales para la demencia (DSM IV y V, ICD-10)

DSM-IV (10)	DSM-V (32)	ICD-10 (33)
A. Desarrollo de deficiencia cognitiva múltiple, manifestada por <ul style="list-style-type: none"> a. Alteración de memoria (aprender nueva información y evocar la ya aprendida), y b. Una o más de las siguientes alteraciones cognitivas: a.- Afasia; b.- Apraxia; c.- Agnosia; d.- Alteración de funciones ejecutivas. 	A. Deterioro cognitivo en relación a un nivel previo en uno o más dominios: <ul style="list-style-type: none"> a. Reporte del paciente o familiar u observación clínica. b. Deterioro cognitivo sustancial encontrados en evaluación objetiva: >2 DE debajo del promedio O menos del tercer percentil. 	La demencia (F00-F03) es un síndrome debido a una enfermedad cerebral, el cual es usualmente de naturaleza crónica o progresiva, en la cual existe un deterioro de múltiples funciones corticales superiores, incluyendo la memoria, el razonamiento, la orientación, la comprensión, el cálculo, el aprendizaje y el lenguaje. La conciencia no está comprometida. El deterioro de la función cognitiva está acompañado, o en algunas ocasiones precedido, por deterioro en el control de las emociones, conducta social o en la motivación. Este síndrome ocurre en la enfermedad de Alzheimer, en enfermedad cerebrovascular, o en otras condiciones que afectan el cerebro de manera primaria o secundaria.
B. Las alteraciones previas representan un deterioro respecto a las capacidades previas del paciente, y producen dificultades significativas en las funciones ocupacional y social.	B. El deterioro cognitivo interfiere con las actividades de la vida diaria.	
C. La evolución se caracteriza por instauración gradual y deterioro cognitivo continuo.	C. No ocurre en el curso de un delirium	
D. Las alteraciones expresadas en A no se deben a lo siguiente: <ul style="list-style-type: none"> a. Otros trastornos del sistema nervioso central que puedan ocasionar deterioro progresivo de la memoria y de otras funciones cognitivas. b. Trastornos sistémicos que pueden ocasionar demencia c. Intoxicaciones 	D. No es debido a otra enfermedad psiquiátrica.	
D. Las alteraciones no ocurren únicamente durante un síndrome confusional agudo.		
E. El trastorno no es atribuible a una alteración psiquiátrica que pudiera justificar las manifestaciones, como por ejemplo una depresión mayor o una esquizofrenia.		

Los criterios diagnósticos para la demencia por EA de 1984 han sido revisados y modificados en el 2011 por el grupo de trabajo de NIA-AA (9). De este modo, fueron actualizados por el NINCDS-ADRDA en criterios clínicos (probable), un curso atípico sin diagnóstico diferencial

(posible) o evidencia neuropatológica (definitivo) (34). El deterioro cognitivo característico de la demencia por EA puede ser detectado y documentado a través de la historia clínica del paciente y la evaluación neuropsicológica que incluye pruebas cognitivas breves o baterías neuropsicológicas extensas (34, 35). En la tabla 2 se presentan los criterios NINCDS/ADRDA DE 1984 Y 2011.

Tabla 2. Criterios diagnósticos NINCDS/ADRDA 1984 y revisados NIA-AA 2011 para la demencia por

EA

NINCDS/ADRDA 1984 (9)	NIA-AA 2011 (34)
<p>EA probable</p> <ul style="list-style-type: none"> a. Síndrome demencial demostrado mediante un cuestionario concreto y confirmado por un test neuropsicológico. b. Comprobación del déficit en dos o más dominios cognitivos. c. Deterioro progresivo en la memoria y otras funciones cognitivas. d. Ausencia de un trastorno a nivel de la conciencia. e. Inicio entre los 40 a 90 años de edad. f. ausencia de enfermedades sistémicas o cerebrales que puedan causar los síntomas. 	<p>EA Probable</p> <p>Criterios clínicos centrales:</p> <ul style="list-style-type: none"> a. Cumplimiento de los criterios generales para demencia. b. Inicio progresivo en meses o años. c. Historial claro de deterioro por reporte u observación. d. El deterioro cognitivo evidente en la historia clínica o al examen sigue uno de los siguiente patrones: A) Presentación amnésica: deterioro en el aprendizaje y en el recuerdo. B) Presentación no-amnésica: alteraciones en el lenguaje, funciones visoespacial y ejecutivas. e. No existe evidencia de historia de enfermedad cerebrovascular relacionada, evidencia de DLB, FTD, u otras comorbilidades concurrentes que pueden ser la causa de la disfunción.
<p>EA posible</p> <ul style="list-style-type: none"> a. Síndrome demencial sin causa aparente, aunque hay variaciones en su inicio, su presentación y curso clínico, con respecto a la EA característica. b. Existencia de otro trastorno cerebral u otro proceso sistémico suficiente para causar la demencia, pero que no se considera que sea su causa. c. Si el enfermo presenta un déficit cognitivo aislado gradualmente y progresivo, sin demostrar otros síntomas. 	<p>EA probable con un nivel mayor de certeza:</p> <ul style="list-style-type: none"> a. EA probable con documentación del deterioro: Evidencia de deterioro progresivo en evaluaciones subsecuentes, basada en reporte de informantes y pruebas cognitivas en el contexto de evaluaciones neuropsicológicas. b. EA probable en portador de una mutación genética: Evidencia de una mutación genética como APP, PSEN1 o 2. <p>EA probable con evidencia de procesos fisiopatológicos de EA</p> <ul style="list-style-type: none"> a. Biomarcadores del depósito de proteína β-amiloide. b. Biomarcadores de lesión o degeneración neuronal.
<p>EA segura</p> <ul style="list-style-type: none"> a. El enfermo ha cumplido todos los criterios diagnósticos de EA probable. b. Existen datos confirmatorios patológicos obtenidos mediante biopsia cerebral o necropsia. 	<p>EA Posible</p> <p>Criterios clínicos centrales:</p> <ul style="list-style-type: none"> a. Curso atípico. b. Presentación etiológica mixta: ACV concomitante, rasgos de DLB, otras enfermedades neurológicas que afecte la cognición.
	<p>EA posible con evidencia de procesos fisiopatológicos de EA</p> <p>Esta categoría es para aquellos que cumplen con criterios de una demencia no-EA, pero que tienen bien evidencia de biomarcadores de proceso fisiopatológico de EA, o bien cumplen con los criterios neuropatológicos de EA.</p>

Por otra parte, para el diagnóstico de la demencia vascular (DV) se cuenta con los criterios definidos NINDS AIREN, publicados por Roman y colaboradores en 1993, los cuales toman en cuenta los tres elementos básicos necesarios para éste diagnóstico: pérdida cognoscitiva, lesiones cerebrales en neuroimágenes, y la exclusión de otras causas de demencia (36) (Tabla 3).

Tabla 3. Criterios diagnósticos de NINDS AIREN 1993 para la demencia vascular

Criterios para VaD (36)
<p>I. Criterios obligatorios:</p> <p>1. Demencia: deterioro respecto al nivel previo de la memoria y al menos otras dos funciones cognitivas, suficiente como para interferir en las actividades diarias (independientemente de lo que interfieran las deficiencias físicas). Se excluyen pacientes con alteración del nivel de conciencia, síndrome confusional agudo, psicosis, afasia o alteración sensitivo-motora que impidan la valoración adecuada de las alteraciones neuropsicológicas. También se excluyen los pacientes con alteraciones sistémicas u otras enfermedades cerebrales (como la enfermedad de Alzheimer) que por sí mismas pudieran explicar las alteraciones cognitivas.</p> <p>2. Enfermedad cerebrovascular, demostrada a través de signos focales compatibles con ACV previo con o sin documentación de ACV previo y evidencia de lesiones vasculares en la neuroimagen –TAC o RM.</p> <p>3. Relación entre los apartados 1 y 2, inferida a partir de una o más de las siguientes circunstancias:</p> <ul style="list-style-type: none"> • Inicio de la demencia en los 3 meses siguientes a un ACV • Deterioro marcado de funciones cognitivas • Progresión fluctuante o escalonada de las alteraciones cognitivas
<p>II. Aspectos compatibles con una demencia vascular probable:</p> <p>a. Alteración de la marcha en fase temprana</p> <p>b. Antecedente de inestabilidad y caídas frecuentes</p> <p>c. Aparición precoz de aumento de la frecuencia de micción, urgencia urinaria u otras alteraciones del control vesical no explicables por un trastorno urológico.</p> <p>d. Parálisis pseudobulbar</p> <p>e. Alteraciones en la personalidad o el estado de ánimo, abulia, depresión, labilidad emocional, y otras alteraciones subcorticales como enlentecimiento psicomotor y alteración de funciones ejecutivas.</p>
<p>III. Aspectos que hacen improbable el diagnóstico de demencia vascular:</p> <p>a. Trastorno temprano de la memoria y deterioro progresivo de la memoria y de otras funciones cognitivas, sin que aparezcan en la neuroimagen lesiones cerebrales focales que lo expliquen</p> <p>b. Ausencia de signos neurológicos focales aparte de las alteraciones cognitivas</p> <p>c. Ausencia de lesiones vasculares en TAC o RM</p>
<p>IV. Criterios de demencia vascular posible:</p> <p>Demencia (según I.1), con signos neurológicos focales, en pacientes en los que no podemos disponer de neuroimagen confirmatoria, o en aquellos que no muestran una relación cronológica congruente entre los ACV y la demencia; también en pacientes con evidencia de enfermedad cerebrovascular, en los que la demencia tiene comienzo insidioso o evolución diferente de la esperada (mesetas prolongadas o mejorías).</p>
<p>V. Criterios de demencia vascular confirmada:</p> <p>a. Criterios clínicos de demencia vascular probable</p> <p>b. Evidencia histopatológica de enfermedad cerebrovascular, obtenida a través de biopsia o autopsia</p> <p>c. Ausencia de ovillos neurofibrilares y placas neuríticas de las esperadas por la edad.</p> <p>d. Ausencia de otras alteraciones clínicas o anatomopatológicas capaces de explicar la demencia</p>

Para el diagnóstico de la DFT, se cuenta con los criterios de Lund-Manchester (37), mientras que para el diagnóstico de la demencia por cuerpos de Lewy, se cuenta con los criterios del Consorcio para el estudio de la demencia por cuerpos de Lewy (28, 38). En el 2007 se desarrollaron los criterios para la definición de la demencia asociada a Enfermedad de Parkinson (39)

Tabla 4. Criterios diagnósticos de Lund-Manchester 1994 para el diagnóstico de DFT

Criterios Lund-Manchester 1994 (40)
<p>I. Criterios esenciales:</p> <p>A. Inicio insidioso y progresión gradual</p> <p>B. Trastorno temprano de la conducta social y de las relaciones interpersonales</p> <p>C. Alteración temprana de la capacidad de autorregulación de la conducta personal</p> <p>D. Superficialidad e indiferencia emocional temprana</p> <p>E. Anosognosia</p>
<p>II. Manifestaciones que apoyan el diagnóstico:</p> <p>A. Comienzo antes de los 65 años</p> <p>B. Agrupación familiar de demencia o antecedente de un trastorno similar en un familiar de primer grado</p> <p>C. Trastorno de conducta: Deterioro en el cuidado y en el aseo personal. Inflexibilidad y rigidez mental. Distractibilidad e impersistencia. Conducta hiperoral y cambios en la conducta alimentaria. Conducta estereotipada y perseverativa. Conducta de utilización.</p> <p>D. Alteraciones del habla y del lenguaje: Alteraciones en la faceta expresiva: Pérdidas de espontaneidad y concreción del lenguaje, habla entrecortada. Lenguaje estereotipado. Ecolalia. Perseveración. Mutismo tardío.</p> <p>E. Signos físicos: Reflejos de desinhibición cortical. Incontinencia urinaria. Acinesia, rigidez, temblor. Presión arterial baja y lábil. Parálisis bulbar, debilidad y atrofia muscular, fasciculaciones (enfermedad de la neurona motora)</p> <p>F. Pruebas complementarias: Exploración neuropsicológica (alteración en las pruebas de las funciones frontales, en ausencia de trastorno marcado de la memoria, el lenguaje y la percepción espacial). Pruebas de neuroimagen cerebral (estructural y/o funcional): atrofia de predominio frontal y/o temporal anterior.</p>
<p>III. Aspectos que excluyen el diagnóstico:</p> <p>A. Aspectos clínicos: Inicio súbito, episodios intercurrentes de deterioro agudo. Traumatismo craneal previo próximo al inicio de las manifestaciones. Trastorno de memoria severo en fases iniciales de la demencia. Desorientación espacial. Mioclonías. Debilidad de tipo corticospinal. Ataxia cerebelosa. Coreoatetosis</p> <p>B. Pruebas complementarias</p> <ul style="list-style-type: none"> • Alteraciones de localización predominantemente postrolándica en las pruebas de neuroimagen estructural o funcional, o lesiones cerebrales multifocales en tomografía computada o resonancia magnética. • Resultados en las pruebas de laboratorio indicativos de afectación cerebral de una alteración inflamatoria o metabólica.

Tabla 5. Criterios diagnósticos de McKeith 2005 para la DL y de Emre 2007 para la demencia asociada a EP.

Criterios para DL (38)	Criterios para PDD (39)
<p>1. Rasgo esencial</p> <ul style="list-style-type: none"> • Demencia definida como deterioro cognitivo progresivo de magnitud suficiente para interferir con la función social o laboral normal. • Alteración de la memoria, que puede no ocurrir necesariamente en los estadios tempranos, pero suele ser evidente con la progresión. • Los déficits en test de atención, función ejecutiva y capacidad viso-espacial pueden ser especialmente marcados. 	<p>1. Rasgo esenciales</p> <p>Diagnóstico de enfermedad de Parkinson conforme a los criterios del Queen Square Brain Bank. Síndrome de demencia de aparición insidiosa y progresión lenta, que se desarrolla en el contexto de una EP establecida, diagnosticada por anamnesis y examen clínico y mental, definido como:</p> <ul style="list-style-type: none"> • Afectación de más de un dominio cognitivo. • Deterioro desde el nivel premórbido. • Intensidad del déficit suficiente como para alterar la vida diaria, independientemente de la que pueda deberse a los síntomas autonómicos o motores.
<p>2. Rasgos centrales (2 para DLB probable, 1 para posible)</p> <ul style="list-style-type: none"> • Fluctuaciones de la función cognitiva con variaciones pronunciadas en atención y alerta. • Alucinaciones visuales recurrentes, típicamente bien formadas y detalladas. • Rasgos espontáneos de parkinsonismo. 	<p>2. Rasgos clínicos asociados</p> <ul style="list-style-type: none"> • Rasgos cognitivos: Atención alterada. Funciones ejecutivas alteradas. Funciones visoespaciales alteradas. Memoria alterada. Funciones esenciales del lenguaje preservadas durante mucho tiempo. • Rasgos conductuales: Apatía. Cambios en la personalidad y el humor. Alucinaciones mayoritariamente visuales. Delirios paranoides. Somnolencia diurna excesiva.
<p>3. Rasgos sugestivos (si se da uno o más en presencia de uno o más rasgos centrales puede hacerse un diagnóstico de DLB probable. En ausencia de rasgos centrales, uno o más permiten el diagnóstico de DLB posible. No se puede diagnosticar DLB probable sólo con rasgos sugestivos)</p> <ul style="list-style-type: none"> • Trastorno de la conducta del sueño REM. • sensibilidad a neurolépticos. • Reducción de la recaptación del transportador de dopamina en ganglios basales demostrada por SPECT o PET. 	<p>3. Rasgos que no excluyen PDD, pero que hacen el diagnóstico dudoso.</p> <ul style="list-style-type: none"> • Coexistencia de cualquier otra anomalía que puede por sí misma originar alteración cognitiva, pero que se juzga que no es la causa de la demencia. • Desconocimiento entre la aparición de los síntomas motores y cognitivos. <p>4. Rasgos que sugieren otras alteraciones como causa del trastorno mental y que imposibilitan el diagnóstico de demencia asociada a la enfermedad de Parkinson.</p> <ul style="list-style-type: none"> • Síntomas cognitivos o conductuales que aparecen aisladamente en el contexto de otras patologías, como: síndrome confusional agudo, depresión mayor. • Rasgos compatibles con criterios NINDS-AIREN de demencia vascular probable.
<p>4. Rasgos de soporte</p> <ul style="list-style-type: none"> • Caídas y síncope a repetición. • Pérdida de conocimiento transitoria e inexplicada. • Disfunción autonómica importante. • Alucinaciones en otras modalidades. • Delirios sistematizados. • Depresión. • Preservación relativa de estructuras temporales mediales en TAC o RM. 	<p>Demencia asociada a la enfermedad de Parkinson probable</p> <p>A. Rasgos esenciales ambos deben estar presentes. B. Rasgos clínicos asociados: A) Perfil típico de déficit cognitivo, incluida la alteración de al menos dos o más de los cuatro dominios cognitivos. B) La presencia de al menos un síntoma conductual apoya el diagnóstico de demencia asociada a la enfermedad de Parkinson probable. La ausencia de síntomas conductuales, sin embargo, no excluye el diagnóstico. C. Ninguno de los rasgos del grupo III. D. Ninguno de los rasgos del grupo IV.</p> <p>Demencia asociada a la enfermedad de Parkinson posible</p> <p>A. Rasgos esenciales : ambos deben estar presentes. B. Rasgos clínicos asociados: A) Perfil atípico de alteración cognitiva en uno o más dominios. B) Los síntomas conductuales pueden estar o no presentes. C. Uno o más rasgos del grupo III. D. Ningún rasgo del grupo IV.</p>

Derivado de los grupos de trabajo del NIA-AA del 2011, también se establecieron los criterios para DCL debido a EA (17). El DCL se divide de acuerdo o no al compromiso de memoria (amnésico o no amnésico) y a la alteración de otros dominios (unidominio o multidominio) (7, 8). En el mismo consenso se aprobó la inclusión de los biomarcadores como línea de investigación para el diagnóstico futuro del DCL debido a EA. En el caso de pacientes con sospecha de DCL, el deterioro en la función cognitiva es fundamental y se recomienda su amplia documentación tanto en la visita inicial como en el seguimiento, con el apoyo de un informante válido (8). Las pruebas de tamizaje de la función cognitiva son numerosas y son usadas principalmente en los escenarios de cuidado primario (15), todas ellas complementadas con cuestionarios diligenciados por informantes. La NIA-AA subraya que los valores de estas pruebas en pacientes con DCL pueden estar entre 1-1.5 desviaciones estándar por debajo de los valores normales para la edad y la escolaridad. (17).

Es de anotar que, en general, el uso de biomarcadores no se requiere como una práctica rutinaria en demencia, exceptuando la DV, donde las neuroimágenes con TAC o RM cerebral son fundamentales para visualizar el tamaño y el área de la lesión vascular. Los biomarcadores que involucran estudios de LCR siguen siendo restringidos como herramientas de investigación teniendo en cuenta su alto costo y que no pueden implementarse de manera rutinaria en todos los países. (11). Algunos estudios sugieren que las neuroimágenes estructurales, como la medición volumétrica del hipocampo, podrían ser útiles para la detección de DCL y su posterior progreso a demencia. Asimismo, las herramientas como el FDG-PET pueden detectar áreas de hipometabolismo cerebral que pueden orientar al diagnóstico preclínico de la demencia por EA u otras causas de deterioro cognitivo (15).

JUSTIFICACION

El diagnóstico de la demencia es complejo y se basa en la acumulación de trabajos y publicaciones con una buena evidencia ante la ausencia de un patrón de oro y de una explicación completa de sus causas (41). El uso de herramientas diagnósticas se ha perfeccionado en los últimos años para poder identificar los pacientes con alto riesgo de desarrollar la enfermedad, y así ofrecer un tratamiento que detenga su progresión o corrija las causas que la originan (42). La evaluación de

la denominada “dimensión biológica de la enfermedad”, más allá de las manifestaciones clínicas de la misma, incluye biomarcadores relacionados con el depósito de amiloide y con daño neuronal (43-47).

Diferentes reportes informan acerca de la implementación de los nuevos criterios más allá de los escenarios de investigación. Bocchetta y cols evaluaron en el 2014 la percepción de los investigadores de 63 centros europeos líderes en investigación en EA, respecto a la utilidad clínica de los biomarcadores, por medio de un cuestionario online (48). A los participantes se les preguntó acerca del uso de dichas pruebas (como RM, FDG-PET o evaluación en LCR), así como el peso del resultado en el diagnóstico y la confianza en dicho resultado fuera de los escenarios de investigación. La respuesta de los autores en 37 centros determinó que 92% de ellos hacían uso de imágenes como apoyo al diagnóstico de DCL, siendo la herramienta más empleada la RM (89%), seguido por el FDG-PET (69%); 39% de los centros incluyen el resultado de estas dos pruebas en los reportes clínicos de los pacientes. Asimismo, 84% de los centros realiza estudio de LCR y 68% los informan. Entre 40-60% de los participantes encontraron que los biomarcadores tuvieron un impacto significativo para determinar el diagnóstico, y se atribuyó la mayor confianza diagnóstica en presencia de un marcador de amiloide más un marcador de daño neuronal (48).

Más recientemente, Molinuevo y colaboradores publicaron un consenso de expertos referente al uso de biomarcadores en LCR para el diagnóstico de demencia por EA, señalando que todos los pacientes con queja subjetiva de memoria y/o aquellos remitidos a evaluación por clínica de memoria deben ser tenidos en cuenta para una punción lumbar con el fin de obtener mediciones de biomarcadores como el β - Amiloide, los niveles de Tau Total y Tau Fosforilado (49). En caso de positividad de los biomarcadores a los niveles establecidos por el laboratorio de referencia, el paciente debe ser considerado como “neuroquímicamente compatible” con el diagnóstico de EA. Otros autores han mostrado más cautela respecto a la utilización indiscriminada de herramientas diagnósticas sin la evidencia suficiente de su utilidad. Le Coteur y colaboradores en el 2013 presentaron sus dudas referentes al diagnóstico de demencia por EA, enfatizando las dificultades asociadas al tamizaje y diagnóstico temprano (50). En particular, los autores señalan las implicaciones del diagnóstico en términos de sobre-diagnóstico, eventos adversos para los pacientes, uso de recursos y capacidad de investigación, todo esto asociado a la falta de certeza

respecto a la efectividad de las intervenciones preventivas, la utilidad diagnóstica de los métodos clínicos actuales y de los biomarcadores y la explotación comercial de la demencia por la industria farmacéutica (50). Asimismo, Gauthier y colaboradores en el 2013 enumeraron los retos éticos relacionados con el diagnóstico y manejo de la demencia causada por EA, enfatizando que en la actualidad el clínico se enfrenta a dilemas éticos relacionados con el diagnóstico de pacientes asintomáticos o con pocos síntomas, sobre todo en cuanto a cómo y cuándo discutir con el paciente el diagnóstico (51). Asimismo, señala que ante la desconocida validez de los biomarcadores biológicos teóricamente asociados a la EA, es difícil confiar en que su presencia pueda predecir de manera adecuada la progresión a demencia por EA, así como promover la estigmatización de los sujetos que reciben un diagnóstico de DCL debido a EA sin tener en la actualidad medidas legales vigentes para estos casos. Estas últimas posturas coinciden con la posición oficial del grupo Cochrane de Demencia y Deterioro Cognitivo, el cual en el 2010 expresó su preocupación referente a la inclusión de nuevas herramientas diagnósticas en los criterios clásicos para la demencia por EA (52). Asimismo, señaló que en la actualidad existen diferentes vacíos relacionados con su uso, como el punto de corte óptimo, la influencia de comorbilidades, diferencias de género y edad, la interacción de las diferentes herramientas y el orden en que deben ser empleadas, entre otros aspectos (52).

En la actualidad, documentos como las guías de práctica clínica y las revisiones sistemáticas de la literatura son reconocidos como referencias claves que orientan a los clínicos a elegir la mejor opción disponible para el manejo de sus pacientes. Las guías de práctica clínica ofrecen a los clínicos, pacientes y administradores de salud, entre otros, recomendaciones para la práctica clínica cotidiana de cuidado de un grupo de pacientes específicos, teniendo en cuenta la mejor evidencia disponible para el tema abordado (53). Sin embargo, si bien la metodología para realizar GPC es ampliamente conocida, se conoce que la calidad de las mismas puede variar y de hecho puede ser mejorada (54). Se sabe qué factores como la calidad en el desarrollo, la independencia editorial y los sistemas de clasificación de la evidencia pueden explicar las diferencias a través de diferentes GPCs (55, 56). Existen estudios previos que realizaron aproximaciones a la evaluación de la calidad y diferencias en contenido de las GPC en demencia, sin embargo dichos estudios no realizaron una evaluación completa de todos los aspectos indispensables (57-60).

Asimismo, las revisiones sistemáticas de la literatura constituyen la recopilación válida de toda la evidencia empírica que responde una pregunta específica respecto al manejo de un paciente (61). En efecto, las revisiones sistemáticas de la exactitud diagnóstica (DTA) de una prueba se enfocan en la capacidad de una herramienta diagnóstica para la detección de la enfermedad, expresada en cifras de sensibilidad y especificidad (62). Este tipo de revisiones presenta información para los clínicos, pacientes, cuidadores y administradores de salud, entre otros, referente a si una determinada prueba diagnóstica puede ser empleada como paso inicial para el proceso de diagnóstico de un paciente (triage), como reemplazo de una prueba existente debido a su superioridad técnica y logística (replacement) o como elemento adicional del proceso diagnóstico para la clasificación o pronóstico de cierto subgrupo de pacientes (add-on) (62). Si bien la metodología para el desarrollo de revisiones diagnósticas aún se encuentra en construcción, ya existen diferentes directrices metodológicas para asegurar la calidad de las mismas (63).

Teniendo en cuenta lo anteriormente expuesto, es indispensable clarificar los siguientes aspectos:

- Es necesario conocer cómo se maneja la evidencia en las GPC que abordan el diagnóstico de demencia, ya que representan uno de los primeros documentos en ser analizados por parte de los clínicos para el manejo de los pacientes. Un aspecto importante para alcanzar este objetivo es evaluar la calidad de las GPC para el diagnóstico de demencia.
- Es necesario evaluar si existen revisiones sistemáticas que permitan recopilar la evidencia de las herramientas diagnósticas empleadas en demencia, así como determinar si su calidad permite que los clínicos confíen en la validez de sus conclusiones.
- Para las diferentes herramientas diagnósticas, incluso aquellas que se han vuelto tradicionales en la práctica médica, es indispensable que se conozca su verdadero valor en el diagnóstico de la demencia y su progresión. La evaluación sistemática de su capacidad diagnóstica debe responder a los retos clínicos de la actualidad en el abordaje de estos pacientes.

PREGUNTA DE INVESTIGACION GENERAL

En este contexto se ha formulado la siguiente pregunta de investigación para este trabajo de tesis, desarrollado como compendio de publicaciones:

¿Cuál es la calidad de la evidencia referente a las herramientas diagnósticas empleadas en la actualidad para el abordaje de los pacientes con demencia por Enfermedad de Alzheimer y otras demencias?

4. OBJETIVOS

OBJETIVO GENERAL 1: Evaluar la calidad de las guías de práctica clínica publicadas a nivel mundial que aborden el diagnóstico de la demencia de tipo Alzheimer y otras demencias.

OBJETIVOS ESPECIFICOS:

- Evaluar si existen diferencias en cuanto a la fuerza de la recomendación, valoración de la calidad de evidencia y contenido, entre las diferentes recomendaciones ofrecidas por las GPC de mayor calidad en relación a la evaluación inicial con pruebas cognitivas breves.
- Evaluar si existen diferencias en cuanto a fuerza de la recomendación, valoración de la calidad de evidencia y contenido, entre las diferentes recomendaciones ofrecidas por las GPC de mayor calidad en relación al uso de evaluación neuropsicológica.
- Evaluar si existen diferencias en cuanto a la fuerza de la recomendación, valoración de la calidad de evidencia y contenido, entre las diferentes recomendaciones ofrecidas por las GPC de mayor calidad en relación al uso de biomarcadores de depósito de amiloide y de daño neuronal.

OBJETIVO GENERAL 2: Evaluar la calidad de las revisiones sistemáticas de características operativas de instrumentos para el diagnóstico de demencia por EA y otras demencias.

OBJETIVOS ESPECIFICOS:

- Evaluar la calidad del reporte de las revisiones sistemáticas incluidas por medio del instrumento PRISMA.
- Evaluar la calidad metodológica de las revisiones sistemáticas incluidas por medio de una versión adaptada del instrumento AMSTAR, adecuado para revisiones DTA.

OBJETIVO GENERAL 3: Determinar la exactitud diagnóstica del MMSE para la detección de sujetos con DCL quienes pueden progresar a demencia, demencia por EA y otras formas de demencia.

OBJETIVOS ESPECIFICOS

- Evaluar la heterogeneidad de las características operativas del MMSE por población y umbrales establecidos, entre otros factores.

5. METODOLOGIA

Como se mencionó al inicio del documento, esta tesis se conforma con tres diferentes estudios, que poseen cada uno su correspondiente metodología para responder a los objetivos previamente planteados.

METODOLOGIA DE LA REVISION DE LA CALIDAD DE LAS GUIAS DE PRÁCTICA CLINICA ENFOCADAS EN EL DIAGNÓSTICO DE DEMENCIA

ESTRATEGIA DE BÚSQUEDA Y SELECCIÓN DE ESTUDIOS

Se realizó una revisión sistemática de GPC enfocadas en el diagnóstico de demencia en sitios desarrolladores de guías (como el Scottish Intercollegiate Guidelines y el National Institute for Clinical Excellence, entre otros), en sitios compiladores de GPC (como el National Guidelines Clearinghouse, la National Library of Guidelines y el sitio web de GIN) y en instituciones relacionadas con el manejo de pacientes con demencia (incluidos sitios web de Ministerios de Latinoamerica). La búsqueda en estos sitios fue complementada con una búsqueda en MEDLINE vía PUBMED por medio de la estrategia de búsqueda que se muestra a continuación.

Tabla 6. Estrategia de Búsqueda en MEDLINE vía PUBMED ARTICULO1

1.	"Dementia"[Mesh] OR dement*[ti] OR senil*[ti] OR Alzheimer*[ti]
2.	"Practice Guideline"[ptyp] OR practice guideline*[tiab] OR guideline*[ti] OR recommendation*[ti] OR "Practice guidelines as topic"[Mesh]
3.	1 AND 2

Se incluyeron GPC en inglés, español o portugués publicadas de Enero del 2005 a Diciembre del 2011, excluyendo versiones previas de la misma GPC, guías enfocadas a salud mental en general o documentos enfocados en posiciones de organizaciones científicas (position statements). También se excluyeron GPC que contenían solo temas de tratamiento de la demencia.

Posterior a la recolección de información, dos autores del estudio seleccionaron por título y resumen las referencias a evaluar en texto completo. Los estudios en texto completo fueron

posteriormente evaluados para definir su inclusión definitiva en el estudio por los mismos autores que realizaron la selección inicial.

EVALUACION DE LA CALIDAD DE LAS GPC

Posterior a la selección de referencias, cuatro autores (dos epidemiólogos y dos neurólogos) evaluaron de manera independiente la calidad de las GPC seleccionadas por medio del instrumento AGREE-II (64), proporcionando puntuaciones en los seis dominios sugeridos a través de 23 ítems: Foco y propósito, participación de los interesados (stakeholders), rigor en el desarrollo, claridad de la presentación, aplicabilidad e independencia editorial. Las puntuaciones por dominio se calcularon acorde con la formula recomendada en el manual de usuario del AGREE-II, y la cual se muestra a continuación:

$$Puntuacion\ por\ dominio = \frac{Suma\ puntuaciones\ obtenidas - Puntuación\ mínima\ (1 \times No\ de\ items \times 4)}{Puntuación\ máxima\ (7 \times No\ de\ ítems \times 4) - Puntuación\ mínima}$$

Para evaluar la concordancia entre los diferentes evaluadores, se estimó el coeficiente de correlación intraclase entre las puntuaciones, acompañado de su correspondiente intervalo de confianza. Se uso SPSS Versión 15.0 para realizar dichas estimaciones.

SELECCION DE GPC Y EVALUACION DE CONTENIDO DE LAS RECOMENDACIONES DIAGNOSTICAS

Posterior a la estimación de la puntuaciones de la GPC por el instrumento AGREE-II, se seleccionaron aquellas GPC con puntuaciones mayores a 60% en el dominio “Rigor en el desarrollo”, esto con el fin de analizar con mayor detalle el contenido de sus recomendaciones referentes al diagnóstico de demencia. Los autores consideraron este punto de corte como suficiente para reflejar un uso mínimo de evidencia científica dentro de la construcción de las recomendaciones, así como un proceso transparente y estandarizado en su desarrollo (65).

Para cada GPC seleccionada en esta etapa, tres autores en conjunto (1 epidemiólogo, 2 neurólogos) extrajeron las recomendaciones relacionadas con los siguientes tres tópicos diagnósticos:

1. Uso de instrumentos cognitivos breves para la valoración inicial de pacientes con sospecha de demencia por EA u otras demencias. Se consideró como instrumento breve toda prueba de lápiz y papel empleada para el tamizaje inicial de estos pacientes.
2. Uso de la evaluación neuropsicológica en pacientes con sospecha de demencia por EA u otras demencias. Se consideró como evaluación neuropsicológica a toda evaluación amplia y estandarizada de la función cognitiva realizada por profesionales entrenados, la cual incluyera dimensiones como la memoria, las funciones ejecutivas y habilidades visoespaciales.
3. Uso de biomarcadores de depósito de β -Amiloide o daño neuronal en pacientes con sospecha de demencia por EA u otras demencias. Se incluyeron todas las recomendaciones referentes a niveles de $A\beta_{42}$ y proteína Tau en líquido cefalorraquídeo, evidencia de depósitos de amiloide en PET-SCAN o FDG-PET y presencia de patrones de atrofia de lóbulos temporal lateral o parietal medial en MRI.

Se extrajo el texto de la recomendación (contenido), así como la valoración de la calidad de evidencia, la fuerza de la recomendación y el tipo de estudios incluidos para soportar la recomendación, en los casos en que esta información estuviese disponible. La información disponible se analizó de manera narrativa y se hizo uso de tablas para ilustrar las diferencias entre las diferentes GPCs.

METODOLOGIA DE LA REVISION DE LA CALIDAD DE REVISIONES SISTEMATICAS DE LA LITERATURA ENFOCADAS EN EL DIAGNÓSTICO DE DEMENCIA

CRITERIOS DE INCLUSION

Los estudios a incluir debían declararse como revisiones sistemáticas de la literatura o haber empleado métodos propios de este tipo de estudios, haber realizado una evaluación de la información incluida en comparación con un estándar de referencia para el diagnóstico de demencia (demencia por EA u otras demencias), así como considerar información de las características operativas de las pruebas (sensibilidad, especificidad u otras propiedades similares). Los pacientes incluidos en la revisión debían ser pacientes adultos mayores de 50 años. Finalmente, las revisiones podían versar en cualquiera de los siguientes desenlaces primarios:

- Detección de demencia por EA, demencia por cualquier causa u otras demencias en pacientes con la enfermedad.
- Conversión de pacientes con DCL a demencia por EA, demencia por cualquier causa u otras demencias.
- Conversión de adultos mayores en envejecimiento normal a demencia o DCL.

METODOS DE BÚSQUEDA DE LAS REVISIONES

Se realizó una búsqueda de revisiones sistemáticas DTA relacionadas con demencia en MEDLINE (vía PUBMED), EMBASE, la Librería Cochrane y la Database of Abstracts of Review of Effects (DARE), publicadas a Diciembre del 2013 y por medio de una estrategia de búsqueda prediseñada y adaptada a cada base de datos seleccionada. La estrategia de búsqueda para MEDLINE puede verse a continuación:

Tabla 7. Estrategia de Búsqueda en MEDLINE vía PUBMED ARTÍCULO 2

1.	"Dementia"[Mesh] OR dement*[ti] OR senil*[ti] OR Alzheimer*[ti] OR AD [tiab]
2.	sensitiv* [Title/Abstract] OR sensitivity and specificity [MeSH Terms] OR diagnos* [Title/Abstract] OR diagnosis [MeSH:noexp] OR diagnostic * [MeSH:noexp] OR diagnosis,differential [MeSH:noexp] OR diagnosis [Subheading:noexp]
3.	1 AND 2
4.	"Meta-Analysis as Topic"[Mesh] OR "Meta-Analysis"[Publication Type] OR systematic review [tiab] OR meta analysis[tiab]
5.	4 AND 3

Como método complementario, se revisaron las referencias de los estudios que cumplían los criterios de inclusión. Se excluyeron todas las referencias relacionadas con resúmenes de congresos, debido a la insuficiente información proporcionada por este tipo de documentos.

RECOLECCION DE DATOS Y ANALISIS

La metodología para la recolección de datos se basó en el capítulo 22 del *Handbook of Systematic Reviews of Interventions*, con algunas modificaciones para adecuarse a las revisiones sistemáticas DTA (61). En relación a la selección de estudios, dos autores de manera independiente evaluaron cada referencia para su inclusión/exclusión. En caso de desacuerdos, se decidió por consenso en adición a la opinión de un tercer autor. En relación a la extracción de datos, dos autores de manera independiente extrajeron los datos básicos de cada referencia seleccionada haciendo uso de un formato previamente diseñado. Los ítems contenidos en el formato referentes a la información general pueden verse en la siguiente tabla:

Tabla 8. Información general – Overview de revisiones DTA sobre demencia

1.	Año de publicación
2.	Tamaño muestral (de estar disponible; en otro caso NA).
3.	Tipo de demencia
4.	Prevalencia agrupada (de estar disponible; en otro caso NA).
5.	Años incluidos en la estrategia de búsqueda (de estar disponible; en otro caso NA).
6.	Tipos de estudios incluidos (de estar disponible; en otro caso NA).
7.	Número de estudios incluidos (de estar disponible; en otro caso NA).
8.	Prueba evaluada
9.	Estándar de referencia
10.	Resultados reportados
11.	Calidad/ evaluación de riesgo de sesgo de los estudios incluidos (de estar disponible; en otro caso NA).
12.	Ítems del QUADAS más afectados (de estar disponible; en otro caso NA).

13. Consecuencias de resultados no concluyentes (de estar disponible; en otro caso NA).
14. Eventos adversos asociados a la prueba (de estar disponible; en otro caso NA).
15. Uso de recursos (de estar disponible; en otro caso NA).
16. Fuente de financiamiento (de estar disponible; en otro caso NA).

De especial atención fue determinar los criterios empleados por los autores para la evaluación de la calidad de los estudios primarios, como por ejemplo el uso del instrumento QUADAS (66, 67).

EVALUACION DE LA CALIDAD METODOLOGICA DE LAS REVISIONES INCLUIDAS

Dos autores evaluaron la calidad metodológica de las revisiones incluidas por medio del instrumento AMSTAR (68). Debido a que el instrumento no fue desarrollado para revisiones DTA, los autores del presente estudio incorporaron definiciones operativas para los 11 ítems que la componen, acordes con las particularidades de la metodología de pruebas diagnósticas y haciendo uso de los capítulos publicados del Cochrane Handbook for DTA reviews (63). En caso de desacuerdos, se decidió por consenso en adición a la opinión de un tercer autor. Estas definiciones operativas pueden verse en la siguiente tabla.

Tabla 9. Definiciones operativas herramienta AMSTAR empleadas en esta revisión

AMSTAR ITEM	Operational definition
1. Was an 'a priori' design provided?	The research question and inclusion criteria should be established before the conduct of the review.
2. Was there duplicate study selection and data extraction?	There should be at least two independent data extractors and a consensus procedure for disagreements should be in place.
3. Was a comprehensive literature search performed?	At least two electronic sources should be searched. The report must include years and databases used (e.g. Central, EMBASE, and MEDLINE). Key words and/or MESH terms must be stated and where feasible the search strategy should be provided. All searches should be supplemented by consulting current contents, reviews, textbooks, specialized registers, or experts in the particular field of study, and by reviewing the references in the studies found. Note: Include MEDLINE, EMBASE, DARE, MEDION, search for grey literature and search on reference list of primary studies.
4. Was the status of publication (i.e. grey literature) used as an inclusion criterion?	The authors should state that they searched for reports regardless of their publication type. The authors should state whether or not they excluded any reports (from the systematic review), based on their publication status, language etc.
5. Was a list of studies (included and excluded) provided?	A list of included and excluded studies should be provided.
6. Were the characteristics of the included studies provided?	In an aggregated form such as a table, data from the original studies should be provided on the participants, interventions and outcomes. The ranges of characteristics in all the studies analyzed e.g. age, race, sex, relevant socioeconomic data, disease status, duration, severity, or other diseases should be reported. Note: Type of studies (cross-sectional, CyC, nested CyC, cohorts, ECC's control group). Spectrum of patients, index test (Add-on, Replacement), prevalence estimated/range, target condition defined, clinical reference standard.

AMSTAR ITEM	Operational definition
7. Was the scientific quality of the included studies assessed and documented?	'A priori' methods of assessment should be provided (e.g., for effectiveness studies if the author(s) chose to include only randomized, double-blind, placebo controlled studies, or allocation concealment as inclusion criteria); for other types of studies alternative items will be relevant. Note: QUADAS I-II, but other checklist could be accepted. STARD and similar doesn't evaluate risk of bias (quality of report). Operational definitions of QUADAS items are desirables.
8. Was the scientific quality of the included studies used appropriately in formulating conclusions?	The results of the methodological rigor and scientific quality should be considered in the analysis and the conclusions of the review, and explicitly stated in formulating recommendations. Note: Include sensitivity analysis or subgroup analysis.
9. Were the methods used to combine the findings of studies appropriate?	Heterogeneity is common on DTA reviews. Consider for meta-analysis a random effect model. Adequate models recommended by Cochrane Collaboration include Bivariate model and Hierarchical SROC. I-square is not recommended in DTA reviews. Sources of heterogeneity on dementia studies: 1. Index test a. Thresholds b. Technical features (including different versions of the test) c. Operator characteristics e.g. training 2. Target disorder a. Reference standard/s used: DSM definition, ICD definition, NINDS-ARDRA, or other classification, including pathological definitions; and operationalisation of these classifications (e.g. individual clinician / algorithm / consensus group). b. Spectrum of target disorder (may also depend on study design) 3. Target population a. Sociodemographic characteristics b. Other characteristics e.g ApoE status, definition and duration of MCI at baseline (if applicable) c. Prevalence in different settings d. Treatment, previous or current interventions 4. Study quality a. Types of studies: see below b. Blinding. Prior clinical information will increase accuracy of the index test. c. Time from index test to reference standard (measured in days or weeks for cross-sectional studies) d. Duration of follow-up (measured in years for delayed-verification studies) e. Loss to follow-up
10. Was the likelihood of publication bias assessed?	Quantitative methods for exploring reporting bias are not well established for studies of DTA. Specifically, funnel plots of the DOR versus the standard error of this estimate will not be considered.
11. Was the conflict of interest stated?	Potential sources of support should be clearly acknowledged in both the systematic review and the included studies.

Asimismo, con el fin de evaluar la calidad en el reporte de las revisiones DTA incluidas, dos autores de manera independiente aplicaron el instrumento PRISMA (69). Los 27 ítems del PRISMA evalúan información que debe incluir toda revisión sistemática, incluyendo aspectos del método empleado, las técnicas estadísticas usadas y el reporte de las fuentes de financiación, entre otros aspectos. En caso de desacuerdos, se decidió por consenso en adición a la opinión de un tercer autor.

SINTESIS DE LA INFORMACION

Se realizó una síntesis narrativa de la información extraída de las revisiones identificadas. Se hizo uso de tablas y figuras para representar la información de las puntuaciones de los instrumentos de calidad aplicados. Para ilustrar las frecuencias de los resultados de información básica, se reportaron valores absolutos y porcentajes para cada caso.

METODOLOGIA DE LA REVISION DE LA CALIDAD DE LOS ESTUDIOS DIAGNÓSTICOS ENFOCADAS EN EL MMSE COMO HERRAMIENTA PARA EVALUAR LA PROGRESIÓN A DEMENCIA EN PERSONAS CON DCL

CRITERIOS PARA LA INCLUSIÓN DE ESTUDIOS

Tipos de estudios: Se incluyeron estudios longitudinales los cuales reportaran resultados del MMSE en pacientes con DCL en una evaluación inicial, y se hubiese evaluado por medio de un patrón de referencia la presencia o no de demencia por EA u otras demencias en un seguimiento de al menos 12 meses. Se excluyeron estudios de corte transversal, estudios antes- después y reportes de caso.

Participantes: Se incluyeron estudios con sujetos diagnosticados con DCL en la evaluación inicial, tanto aquellos derivados de la comunidad como aquellos derivados de escenarios de cuidado primario o especializado. El diagnóstico de DCL debía haber sido establecido por medio de los Criterios de Petersen (70, 71), los criterios de Matthews (72), o por medio de las puntuaciones del Clinical Dementia Rating ($= 0.5$)(73). Los anteriores criterios incluyen queja subjetiva del paciente, declive de la memoria verificado de manera objetiva por tests neuropsicológicos en combinación con la historia clínica del paciente, declive en otros dominios cognitivos, así como deterioro mínimo o inexistente en las actividades cotidianas. Se incluyeron todos los subtipos de DCL (amnésico dominio único, amnésico dominio múltiple, no- amnésico dominio único, no- amnésico dominio múltiple). Se excluyeron estudios en los que los participantes hubiesen desarrollado deterioro cognitivo por causas como abuso de alcohol o sustancias alucinógenas, traumas

asociados con el sistema nervioso central (por ejemplo, hemotomas subdurales), tumores o infecciones, entre otros.

Prueba a evaluar: La prueba Mini-Mental o MMSE es una prueba breve tipo lápiz y papel de la función cognitiva, basada en una puntuación máxima de 30 unidades, la cual incluye ítems que evalúan la orientación, la concentración, la atención, la memoria verbal, la denominación y las habilidades viso-espaciales (74). En estudios de tipo longitudinal, los pacientes con DCL son evaluados con el MMSE para obtener una puntuación al inicio del estudio, y entonces son seguidos por un número determinado de meses para identificar los nuevos casos de demencia. Su utilidad para predecir la aparición de la demencia puede ser evaluada en diferentes puntos de corte, algunos de ellos previamente especificados u otros estimados por medio de métodos estadísticos (por ejemplo, por medio de regresión logística). Los puntos de corte óptimos se establecen acorde con los valores de sensibilidad y especificidad, entre otros criterios.

Condición a evaluar: La condición a evaluar es la progresión de DCL a demencia por cualquier causa, demencia asociada a AD u otras formas de demencia, luego de un seguimiento de al menos 12 meses.

Estándar de referencia: En la actualidad no existe un patrón de oro plenamente establecido para el diagnóstico de demencia, e incluso los diagnósticos basados en criterios neuropatológicos han sido cuestionados (31). En la presente revisión se emplearon patrones de referencia comúnmente aceptados por diferentes estudios, basados en el diagnóstico clínico. Estos criterios diagnósticos incluyen los establecidos por el Manual Diagnóstico y Estadístico de Enfermedades Mentales (DSM-IV), así como aquellos presentados por la Clasificación Internacional de Enfermedades (ICD-10). De igual manera, los criterios establecidos por el Instituto Nacional de Enfermedades Neurológicas, de la Comunicación y Accidentes Cerebrovasculares (NINCDS), junto con la Asociación para la Enfermedad de Alzheimer y sus desórdenes asociados (ADRDA), son los criterios con mayor aceptación en la actualidad para la definición de la demencia por AD (9, 34). Los criterios de McKeith (28, 38) son los más ampliamente aceptados para el diagnóstico de demencia por cuerpos de Lewy, mientras que para las demencias frontotemporal y Vascular se cuenta con los criterios de Lund-Manchester (37) y los criterios del Instituto Nacional de Enfermedades

Neurológicas, de la Comunicación y Accidentes Cerebrovasculares (NINCDS), junto con la Asociación Internacional para la Investigación y enseñanza en Neurociencias (AIREN) (36), respectivamente.

MÉTODOS DE BÚSQUEDA

Búsquedas electrónicas: Se buscaron posibles referencias en MEDLINE (vía OVID, 1966 a Mayo 2014), EMBASE (vía OVID, 1982 a Mayo 2014), BIOSIS (vía OVID, desde el inicio a Mayo 2014), Science Citation Index (ISIWeb of Knowledge, desde su inicio hasta Mayo 2014), PsycINFO (vía OVID, desde el inicio a Mayo 2014) y LILACS (vía BIREME, 1982 a Mayo 2014). La base de Science Citation Index incluye resúmenes de congresos. La estrategia de búsqueda diseñada para MEDLINE puede verse a continuación:

Tabla 10. Estrategia de búsqueda- MEDLINE (Vía OVID. 1946- 2014)

1. MMSE*.ti,ab.
2. sMMSE.ti,ab.
3. Folstein*.ti,ab.
4. MiniMental.ti,ab.
5. "mini mental stat*".ti,ab.
6. or/1-5

Se diseñaron estrategias de búsqueda similares para cada una de las bases previamente mencionadas, haciendo uso de vocabulario avanzado según el caso (términos MeSH o Emtree). No se emplearon filtros para limitar el número de referencias para tamizar, debido a que su uso no ha probado ser sensible para la identificación de estudios diagnósticos (75). Asimismo, no se aplicaron restricciones de idioma a ninguna de las búsquedas realizadas. De manera complementaria, se realizaron búsquedas en el registro Cochrane de estudios diagnósticos (desarrollado por el grupo Cochrane de Enfermedad Renal), así como en el registro especializado del grupo de Demencia de Cochrane (ALOIS), el cual incluye estudios de intervenciones y de exactitud diagnóstica. El coordinador de búsquedas del grupo de Demencia desarrolló todas las búsquedas y proporcionó los resultados finales al grupo de la revisión.

Búsqueda en otras fuentes: Se revisaron las referencias de todos los estudios potenciales en búsqueda de estudios adicionales. Asimismo, se buscó información adicional en:

- MEDION (Meta-analyses van Diagnostisch Onderzoek): www.mediondatabase.nl.
- DARE (Database of Abstracts of Reviews of Effects): www.york.ac.uk/inst/crd/crddatabases.html.
- HTA Database (Health Technology Assessments Database, The Cochrane Library).
- ARIF database (Aggressive Research Intelligence Facility): www.arif.bham.ac.uk.

A través de PUBMED se usó la herramienta de “estudios relacionados” para identificar estudios relevantes no identificados previamente por ninguna de las estrategias previamente descritas. Asimismo, se identificó la literatura gris a través del Science Citation Index, el cual contiene información acerca de resúmenes presentados en congresos. No se realizó búsqueda manual en revistas especializadas en el tema, debido a que existe reducida evidencia respecto a los beneficios de esta técnica en relación a estudios de exactitud diagnóstica (76).

RECOLECCION DE DATOS Y ANALISIS

Selección de estudios: En primer lugar, se seleccionaron las referencias derivadas de las estrategias de búsqueda por medio de sus títulos y resúmenes, labor que realizaron los autores de la revisión junto con un equipo amplio de autores de otras revisiones relacionadas. Luego de la selección inicial, se obtuvo el texto completo de todas las referencias potencialmente elegibles para revisión de dos autores de la revisión, quienes de manera independiente evaluaron cada referencia frente a los criterios de inclusión y exclusión. En caso de desacuerdo se resolvieron las discrepancias por medio de una discusión amplia de la referencia, o en caso de no llegar a acuerdo, se incluyó la referencia dentro del estudio. Todos los resultados de este proceso fueron ilustrados en un diagrama PRISMA de información.

Selección de estudios: La información referente a las características de los estudios incluidos fue extraída en un formato previamente diseñado para la presente revisión, el cual incluyó toda la información relacionada con la evaluación de la calidad y las fuentes de heterogeneidad. Dicho formato fue probado con cinco estudios diagnósticos previo a su uso extendido en la presente revisión. La extracción de estos datos fue realizada por dos autores. De igual manera se

identificaron los datos numéricos referentes a los verdaderos y falsos positivos o negativos, los cuales fueron representados en tablas de RevMan.

Evaluación de la calidad metodológica: Se evaluó la calidad metodológica de cada uno de los estudios incluidos por medio de la herramienta QUADAS-II (67), la cual es la escala recomendada por la Colaboración Cochrane para dicha evaluación. Cada uno de los dominios de esta escala (Selección de pacientes, prueba evaluada, estándar de referencia y flujo de pacientes) fue valorado en términos del riesgo de sesgo, y en los tres primeros dominios también se valoró si habían cuestiones de aplicabilidad. Se desarrollaron definiciones operativas para cada uno de los dominios de la herramienta, los cuales se muestran en la siguiente tabla.

Tabla 11. Definiciones operativas instrumento QUADAS-II

Domain 1: patient selection	
Risk of bias: could the selection of patients have introduced bias? (high, low, unclear)	
<i>Was a consecutive or random sample of patients enrolled?</i>	When sampling is used, the methods least likely to cause bias are consecutive sampling and random sampling, which should be stated and/or described. Non-random sampling or sampling based on volunteers is more likely to be at high risk of bias.
<i>Was a case-control design avoided?</i>	Case-control study designs have a high risk of bias, but sometimes they are the only studies available, especially if the index test is expensive and/or invasive. Nested case-control designs (systematically selected from a defined population cohort) are less prone to bias, but they will still narrow the spectrum of participants who receive the index test. Other study designs (both cohort and case-control) that may increase bias are those designs for which the study team deliberately increases or decreases the proportion of participants with the target condition, for example, a population study may be enriched with extra participants with dementia from a secondary care setting.
<i>Did the study avoid inappropriate exclusions?</i>	The study will be automatically graded as unclear if exclusions are not detailed (pending contact with study authors). When exclusions are detailed, the study will be graded as "low risk" if exclusions are believed by the review authors to be appropriate. Exclusions common to many studies of dementia include the following: medical instability; terminal disease; alcohol/substance misuse; concomitant psychiatric diagnosis; and other neurodegenerative condition. However, if "difficult to diagnose" groups are excluded, this may introduce bias, so exclusion criteria must be justified. For a community sample, we would expect relatively little exclusion. Post hoc exclusions will be labelled "high risk" of bias.
Applicability: are there concerns that included patients do not match the review question? (high, low, unclear)	Included patients should match the intended population as described in the review question. If not already specified in the review inclusion criteria, the setting will be particularly important- the review authors should consider population in terms of symptoms, pretesting and potential disease prevalence. Studies that use very selected participants or subgroups will be classified as having low applicability, unless they are intended to represent a defined target population, for example, people with memory problems referred to a specialist and investigated by lumbar puncture.
Domain 2: index test	
Risk of bias: could the conduct or interpretation of the index test have introduced bias? (high, low, unclear)	
<i>Were MMSE results interpreted without knowledge of the reference standard?</i>	Terms such as "blinded" or "independently and without knowledge of" are sufficient, and full details of the blinding procedure are not required. This item may be scored as "low risk" if explicitly described, or if a clear temporal pattern to the order of testing precludes the need for formal blinding (e.g. all MMSE assessments were performed before the dementia assessment). As most neuropsychological tests are administered by a third party, knowledge of the dementia diagnosis may influence ratings; tests that are self-administered, for example, use of a computerized version, may be associated with less risk of bias.
<i>Were MMSE thresholds prespecified?</i>	For neuropsychological scales, there is usually a threshold above which participants are

	classified as “test positive”; this may be referred to as the <i>threshold</i> , the <i>clinical cut-off</i> or the <i>dichotomisation point</i> . Different thresholds are used in different populations. A study is classified as having higher risk of bias if the study authors define the optimal cut-off post hoc on the basis of their own study data. Some papers use an alternative methodology for analysis that does not use thresholds; these papers should be classified as not applicable.
<i>Were sufficient data on MMSE application given for the test to be repeated in an independent study?</i>	Particular points of interest includemethod of administration (e.g. self- completed questionnaire vs direct questioning interview); nature of the informant; and language of the assessment. If a novel form of the index test is used, for example, a translated questionnaire, details of the scale should be included and a reference given to an appropriate descriptive text; evidence of validation should be provided.
Applicability: are there concerns that the index test, its conduct or its interpretation may differ from the review question?	Variations in length, structure, language and/or administration of the index test may affect applicability if they vary from those specified in the review question.
Domain 3: reference standard	
Risk of bias: could the reference standard, its conduct or its interpretation have introduced bias? (high, low, unclear)	
<i>Is the assessment used for clinical diagnosis of dementia acceptable?</i>	Commonly used international criteria to assist with clinical diagnosis of dementia include those detailed in <i>DSM-IV</i> and <i>ICD-10</i> . Criteria specific to dementia subtypes include but are not limited to NINCDS-ADRDA criteria for Alzheimer’s dementia, McKeith criteria for Lewy body dementia, Lund-Manchester criteria for fronto-temporal dementia and NINDS-AIREN criteria for vascular dementia. When the criteria used for assessment are not familiar to the review authors and the Cochrane Dementia and Cognitive Improvement Group, this item should be classified as “high risk of bias”.
<i>Was clinical assessment for dementia performed without knowledge of the MMSE results?</i>	Terms such as “blinded” and “independent” are sufficient, and full details of the blinding procedure are not required. Interpretation of results of the reference standard may be nuanced by knowledge of results of the index test.
Applicability: are there concerns that the target condition as defined by the reference standard does not match the review question? (high, low, unclear)	Some methods of dementia assessment, although valid, may diagnose a far smaller or larger proportion of individuals with the disease than in usual clinical practise. For example, currently the reference standard for vascular dementia may underdiagnose compared with usual clinical practise. In this instance, the item should be rated as having poor applicability.
Domain 4: participant flow and timing	
Risk of bias: could the participant flow have introduced bias? (high, low, unclear)	
<i>Was there an appropriate interval between MMSE and the reference standard?</i>	As we test the accuracy of the MMSE test for MCI conversion to dementia, a delay will always be noted between the index test and the reference standard assessments. The time between reference standard and index test will influence the accuracy, and therefore we will note time as a separate variable (both within and between studies) and will test its influence on diagnostic accuracy. We have set a minimum mean time to follow-up assessment of 1 year. If more than 16% of participants undergo assessment for MCI conversion before 9 months, this item will score “no”.
<i>Did all participants receive the same reference standard?</i>	In some scenarios, participants who score “test positive” on the index test may undergo a more detailed assessment for the target condition. When dementia assessment (or reference standard) differs between participants, this should be classified as high risk of bias.
<i>Were all participants included in the final analysis?</i>	If the number of participants enrolled differs from the number of participants included in the 2 × 2 table, the potential for bias exists. If participants lost to follow-up differ systematically from those who remain, then estimates of test performance may differ. If dropouts are present, these should be accounted for; the maximum proportion of drop-outs for low risk of bias has been specified as 20%. Details of the causes of study drop-outs are crucial, and if such data are missing, the reliability of the conclusions must be questioned.

Evaluación de la calidad metodológica: La condición a estudio incluyó dos categorías: 1) la demencia en general- por cualquier causa; 2) subtipos de demencia (por EA, DV, etc). Los estudios incluidos podrían tener detalles para más de una condición. Para el análisis de datos, dicha información se manejó separadamente, haciendo uso de la información de verdaderos y falsos positivos/negativos para el cálculo de la sensibilidad y especificidad individual, con sus respectivos

intervalos de confianza del 95%. Los estudios incluidos se presentaron de manera gráfica por medio de su ubicación en el espacio ROC de sus características operativas individuales, haciendo uso del programa RevMan para este análisis descriptivo. Cuando un determinado estudio presentaba más de un punto de corte, estos fueron representados en la misma gráfica.

Para realizar las estimaciones meta-analíticas (que en el caso de la presente revisión no se desarrollaron para obtener estimadores conjuntos de las características operativas, sino para obtener los parámetros de la curva estimada de descripción de los datos), se usó el comando Metandi de STATA 13 (StataCorp, College Station, Texas), el cual estima los parámetros de un modelo bivariado de efectos aleatorios (77). Ya que no se incluyeron covariables al modelo, este procedimiento permite realizar una aproximación a los resultados del modelo jerárquico HSROC (78). Los datos del modelo se emplearon para graficar la curva ROC de cada conjunto de datos en REvMan. Debido a que se decidió no estimar las características operativas conjuntas debido a la heterogeneidad de los datos, se calculó la especificidad mediana del conjunto de datos, y bajo los parámetros del modelo se obtuvo la sensibilidad estimada; esto permitió el cálculo de una aproximación de la exactitud diagnóstica de la prueba en una cohorte hipotética de 100 pacientes con DCL. Esta aproximación descriptiva de la información, en ausencia de la posibilidad de un metanálisis formal, ya ha sido usada en otras revisiones diagnósticas.

Investigación de la heterogeneidad: Se examinaron de manera visual las gráficas que ilustraban los datos de sensibilidad y especificidad de los estudios incluidos, con el fin de explorar si había signos de heterogeneidad en los datos ilustrados. Si bien se había planeado hacer una evaluación formal de diferentes fuentes de heterogeneidad por el ajuste de modelos HSRROC, la presencia de heterogeneidad en la inspección inicial, junto con el insuficiente número de estudios identificados, no permitieron la realización de dicho plan de análisis.

Análisis de sensibilidad: Debido al limitado número de estudios identificados para cada conjunto de datos, sólo se realizaron análisis de sensibilidad respecto a la versión del MMSE empleada en el diagnóstico de demencia por EA.

Evaluación del sesgo de reporte: En la actualidad no existen métodos cuantitativos para explorar el efecto del sesgo de reporte en revisiones diagnósticas. Específicamente, los autores de la presente

revisión no consideraron emplear la evaluación de OR diagnósticos en comparación con su Error estándar.

4. RESULTADOS

En esta sección se presentan los resultados alcanzados para cada uno de los objetivos principales previamente descritos en este documento.

Alzheimer's Disease Dementia Guidelines for Diagnostic Testing: A Systematic Review

Ingrid Arevalo-Rodriguez, Olga L. Pedraza, Andrea Rodríguez, Erick Sánchez, Ignasi Gich, Ivan Solà, Xavier Bonfill and Pablo Alonso-Coello.

American Journal of Alzheimer's Disease & Other Dementias (2012): 28(2), 111-119

Factor de impacto: 1.426.

RESUMEN DE RESULTADOS

Se recopilaron un total de 537 referencias, excluyendo posteriormente 518 que no estuvieron relacionadas con demencia por EA. Al final se contó con 15 GPC referentes al diagnóstico de demencia por EA distribuidas en 22 referencias (2, 35, 79-98). En el análisis por medio del instrumento AGREE-II se encontró que los dominios con mayores puntuaciones fueron los de "Foco y propósito" y "Claridad de presentación" (75.4% y 80%, respectivamente), mientras los dominios de "Aplicabilidad" e "Independencia Editorial" obtuvieron las menores puntuaciones (32.4% y 47%, respectivamente). Sólo 1 GPC obtuvo puntuaciones mayores a 60% en todos los dominios (81). 40% de las GPC analizadas obtuvieron puntuaciones de 60% o superiores en el dominio de "Rigor en el desarrollo" (2, 35, 81, 82, 92, 97), siendo estas GPC clasificadas como de alta calidad.

Respecto a la clasificación de la evidencia, 4 de las 15 GPC incluidas no proporcionaron un reportes de un sistema para graduar la evidencia o las recomendaciones. Las 6 GPC de alta calidad identificadas previamente hacen uso de 4 diferentes sistemas de graduación, siendo el más común el desarrollado por el Scottish Intercollegiate Guidelines Network (66%) (99). Solo una de la GPC empleó un sistema diseñado para el abordaje de preguntas de naturaleza diagnóstica(92).

Todas las GPC seleccionadas abordan al menos un tópico diagnóstico de los evaluados en esta revisión. 93% de las GPC proporcionan recomendaciones referentes al uso de pruebas cognitivas

breves en el diagnóstico de pacientes con sospecha de demencia por EA, 60% presentan recomendaciones referentes al uso de pruebas neuropsicológicas y 26% GPC presentan recomendaciones respecto al uso de biomarcadores de daño neuronal y acumulación de β -Amiloide.

Recomendaciones referentes a las pruebas cognitivas breves

Seleccionando sólo las GPC clasificadas como de alta calidad, todas ellas proporcionan recomendaciones referentes al uso de pruebas cognitivas breves en el diagnóstico de demencia por EA. 50% de las GPC no identifican una prueba en particular, pero todas ellas insisten en el uso de instrumentos objetivos y válidos para dicha evaluación cognitiva. En las restantes GPC, la prueba más recomendada fue el MMSE, acompañada por un total de 19 instrumentos cognitivos como el 7- minute Screen y el Montreal Cognitive Test (MOCA). Cuatro de las 6 GPC de alta calidad (83%) consideran información de pruebas validadas y estandarizadas en el contexto local, con dos de ellas recomendando el MMSE en una versión validada en el país de desarrollo de la guía.

En relación al tipo de estudio incluido, los estudios tipo cohorte o corte-transversal fueron los más frecuentes. Al revisar las recomendaciones se encontró un considerable desacuerdo en relación a la valoración de la calidad de la evidencia y la confianza en la recomendación. La calidad de la evidencia fue clasificada desde la más baja calidad hasta la más alta, dependiendo esta clasificación del sistema empleado por la GPC. De igual manera, la fuerza de las recomendaciones osciló desde niveles altos (A) hasta niveles intermedio-bajo (C).

Recomendaciones referentes a las pruebas neuropsicológicas

Todas las GPC de alta calidad presentaron recomendaciones referentes al uso de pruebas neuropsicológicas, en el contexto de dificultades en el diagnóstico relacionadas con el estadio de la demencia o para el diagnóstico diferencial. Los estudios de tipo observacional fueron los diseños más frecuentes en la evidencia identificada. Nuevamente, se encontró gran discrepancia en la manera de valorar la calidad de la evidencia (Desde muy baja a Alta calidad) y la fuerza de la recomendación (Desde niveles A hasta D). Ninguna de las GPC propuso una batería

neuropsicológica específica, y sólo una GPC proporcionó dentro de la recomendación detalles referentes a los dominios que debían ser evaluados por dichas pruebas (92).

Recomendaciones referentes al uso de biomarcadores

66% de las GPC de alta calidad presentaron recomendaciones explícitas referentes a biomarcadores de acumulación de amiloide y de daño neuronal. Tres de las seis GPC seleccionadas enfatizaron la falta de estudios clínicos para estas herramientas diagnósticas y proporcionaron recomendaciones en contra de su uso (2, 82, 97). En un cuarto caso, se presenta una recomendación no clara (a favor/ en contra) respecto al uso del FDG-PET y los niveles de Tau en LCR (92). Nuevamente, los estudios observacionales son la evidencia reportada con mayor frecuencia para apoyar dichas recomendaciones, y se encontró una similar variación de la calificación de la calidad de la evidencia y en la fuerza de la recomendación (Niveles B a C) que la previamente expuesta en el resto de recomendaciones.

Alzheimer's Disease Dementia Guidelines for Diagnostic Testing: A Systematic Review

American Journal of Alzheimer's Disease & Other Dementias®
28(2) 111-119
© The Author(s) 2012
Reprints and permission:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/1533317512470209
aja.sagepub.com



Ingrid Arevalo-Rodriguez, MSc^{1,6}, Olga L. Pedraza, MD, MSc²,
Andrea Rodríguez, MD, MSc¹, Erick Sánchez, MD²,
Ignasi Gich, MSc³, Ivan Solà, MSc³, Xavier Bonfill, MD, PhD^{4,5}, and
Pablo Alonso-Coello, MD, PhD^{4,5}

Abstract

Alzheimer's disease dementia (AD dementia) is one of the most common neurodegenerative diseases worldwide, with a growing incidence during the last decades. Clinical diagnosis of cognitive impairment and presence of AD biomarkers have become important issues for early and adequate treatment. We performed a systematic literature search and quality appraisal of AD dementia guidelines, published between 2005 and 2011, which contained diagnostic recommendations on AD dementia. We also analyzed diagnostic recommendations related to the use of brief cognitive tests, neuropsychological evaluation, and AD biomarkers. Of the 537 retrieved references, 15 met the selection criteria. We found that Appraisal of Guidelines Research and Evaluation (AGREE)-II domains such as applicability and editorial independence had the lowest scores. The wide variability on assessment of quality of evidence and strength of recommendations were the main concerns identified regarding diagnostic testing. Although the appropriate methodology for clinical practice guideline development is well known, the quality of diagnostic AD dementia guidelines can be significantly improved.

Keywords

diagnosis, Alzheimer's disease dementia, systematic review, clinical practice guidelines

Background

The world population is rapidly aging, which causes a significant increase in chronic and neurodegenerative diseases, with serious consequences in terms of global public health. The prevalence of diseases such as dementia is nearly 42 million patients in 2012, with approximately 4.6 million new cases a year.¹ These figures would increase up to 300% by 2040, with an impact on expenditure of over US\$ 422 billion.^{1,2}

Alzheimer's disease dementia (AD dementia), the most common dementia in the Western world, is defined as a cognitive decline documented by standardized testing on 2 or more domains, which interferes with daily function and represents a decline in previous levels.³ Recently, the National Institute of Neurological and Communicative Disorders and Stroke–Alzheimer's Disease and Related Disorders Association criteria have been updated and presented 3 different types of AD dementia according to presence of core clinical criteria only (probable), an atypical course without differential diagnosis (possible), or neuropathological evidence (definitive).⁴

The presence of cognitive impairment, especially in the memory domain, is a fundamental part of the clinical criteria used to diagnose AD dementia. This impairment must be detected and

diagnosed through a combination of clinical history analysis and objective cognitive assessment by means of a brief mental evaluation or comprehensive neuropsychological testing.⁴ The latter reflects the clinical dimension of the disease and characterizes the type of presentation. However, new trends in AD dementia

¹ Grupo de Evaluación de Tecnologías y Políticas en Salud - GETS, Clinical Research Institute - School of Medicine, National University of Colombia, Bogotá DC, Colombia

² Department of Neurosciences and Memory Clinic Unit, Fundación Universitaria de Ciencias de la Salud, Bogotá, Colombia

³ Iberoamerican Cochrane Centre, Clinical Epidemiology and Public Health Department, Institute of Biomedical Research (IIB Sant Pau), Spain

⁴ Iberoamerican Cochrane Centre, Institute of Biomedical Research (IIB Sant Pau), CIBER Epidemiología y Salud Pública (CIBERESP), Spain

⁵ Universitat Autònoma de Barcelona, Barcelona, Spain

⁶ Doctoral candidate at the Pediatrics, Obstetrics & Gynecology and Preventive Medicine Department, Universitat Autònoma de Barcelona, Barcelona, Spain

Corresponding Author:

Ingrid Arevalo-Rodriguez, MSc, Grupo de Evaluación de Tecnologías y Políticas en Salud - GETS, Clinical Research Institute, School of Medicine, National University of Colombia, Bogotá DC, Colombia.
Email: ingrid.arevalo@e-campus.uab.cat

diagnosis suggest the need to reflect upon the biological dimension of the disease, including biomarkers related to brain amyloid- β protein deposition and downstream neuronal degeneration.⁵

Several clinical practice guidelines (CPGs) for the management of AD dementia have been published in different countries. These offer users (professionals and health administrators, patients and caregivers, among others) clinical recommendations for daily practice based on the best available evidence.⁶ Although CPGs have become a useful clinical tool, their quality can vary and could be improved.⁷ The CPG assessment must be conducted in an objective manner using standardized instruments that address issues such as rigor of development, clarity of recommendations, and applicability, among others.⁸ It is also important to analyze the differences in recommendations among guidelines focused on the same topic, which can vary and even be contradictory.^{9,10}

There are previous studies related to the assessment of the quality and content of AD dementia CPGs, but these were not conducted systematically and did not assess recommendations on AD dementia diagnosis.¹¹⁻¹⁴ The purpose of this review was to assess the quality of CPGs that focus on the diagnosis of AD dementia as well as to examine the recommendations on 3 critical aspects of AD diagnosis: initial testing with brief cognitive screening tools, use of neuropsychological assessment, and use of biomarkers of amyloid- β accumulation and neuronal injury.

Methods

We performed a systematic review of AD dementia CPGs retrieved from the Web sites of guideline developers (Scottish Intercollegiate Guidelines, National Institute for Clinical Excellence, among others), CPG compilers (National Guidelines Clearinghouse, National Library of Guidelines, GIN Web site, among others), and institutions that conduct research related to dementia.

We complemented this search in MEDLINE, using the terms “dementia,” “Alzheimer,” and “elderly,” with a methodological filter to search for CPGs (the full search strategy is available upon request). The CPGs with diagnostic recommendations in English, Spanish, and Portuguese published between January 2005 and December 2011 were included. Previous versions of the same guideline, letters to the editor, or guidelines with a focus on population screening or general mental health were excluded.

Two authors independently selected all potentially eligible studies by reading the titles and abstracts of the retrieved references. Subsequently, 4 trained authors independently assessed the quality of each CPG using the Appraisal of Guidelines Research and Evaluation (AGREE)-II tool.⁸ This instrument includes the evaluation of 6 domains: Scope and Purpose, Stakeholder Involvement, Rigor of Development, Clarity of Presentation, Applicability, and Editorial Independence. The intraclass correlation coefficient (ICC) and its 95% confidence interval (95% CI) were calculated to assess interrater agreement. Statistical analyses were performed using SPSS® 15.0 (SPSS Inc, Chicago, IL).

For the analyses of recommendations regarding diagnostic testing, we only included guidelines of high quality, defined

as those with scores equal to or greater than 60% in the “Rigor of development” domain. We used this cutoff point as an adequate figure to reflect that a valid and transparent process had been adopted in the development of recommendations.¹⁵ For each guideline, 3 authors extracted recommendations concerning: (1) the brief cognitive instruments recommended for initial evaluation of patients with suspected AD dementia, (2) the use of neuropsychological assessment, and (3) the use of biomarkers in patients with suspected AD dementia. The recommendations (sentences) of each CPG were extracted along with the quality of evidence, strength of the recommendation, and type of studies included. This information was synthesized in a recommendation matrix for further analysis. All disagreements were resolved by consensus.

All paper and pencil tests used for the initial assessment of patients with suspected AD dementia were classified as brief cognitive instruments. Neuropsychological assessment was defined as all comprehensive and standardized cognitive evaluations, including appraisal of memory, language, executive functions, and visual-spatial abilities conducted by trained clinicians. Recommendations on the following AD biomarkers were included: low cerebrospinal fluid (CSF) amyloid β_{42} levels, positive retention of tracer in positron emission tomography (PET), elevated Tau protein levels in CSF (total or phosphorylated), decreased fluorodeoxyglucose (FDG) uptake on PET reflected in the temporoparietal cortex, and presence of patterns of atrophy on magnetic resonance imaging (MRI), involving medial, basal and lateral temporal lobes and medial parietal cortex.⁴

Results

We retrieved a total of 537 citations and excluded 518 that were not related to AD dementia. We also excluded letters to the editor as well as CPGs in languages other than English, Spanish, and Portuguese, those with recommendations concerning only a subset of patients with dementia and those with general recommendations for mental health. In all, 4 CPGs containing therapeutic recommendations only were also excluded (Figure 1). We assessed 15 CPGs distributed among 22 individual references.¹⁶⁻³⁷

We found that the domains with the highest scores were “Scope and Purpose” and “Clarity of Presentation” (75.4% and 80%, respectively), whereas “Applicability” and “Editorial Independence” had the lowest figures (32.4% and 47%, respectively). Only 1 (6.7%) CPG had scores above 60% in all domains.¹⁸ All the 6 (40%) guidelines had ratings equal to or above 60% for the “Rigor of Development” domain^{18,19,25,30,34,36} (Table 1). We classified these as guidelines of high quality. Agreement among reviewers was considerable (ICC = 0.78, 95% CI: 0.60-0.91).

Of the 15 identified guidelines, 4 did not report a system to grade evidence or recommendations. Conversely, the 6 high-quality CPGs used 4 different grading systems. The most commonly used system, out of these 4, was that developed by the Scottish Intercollegiate Guidelines Network (66%).³⁸ Only

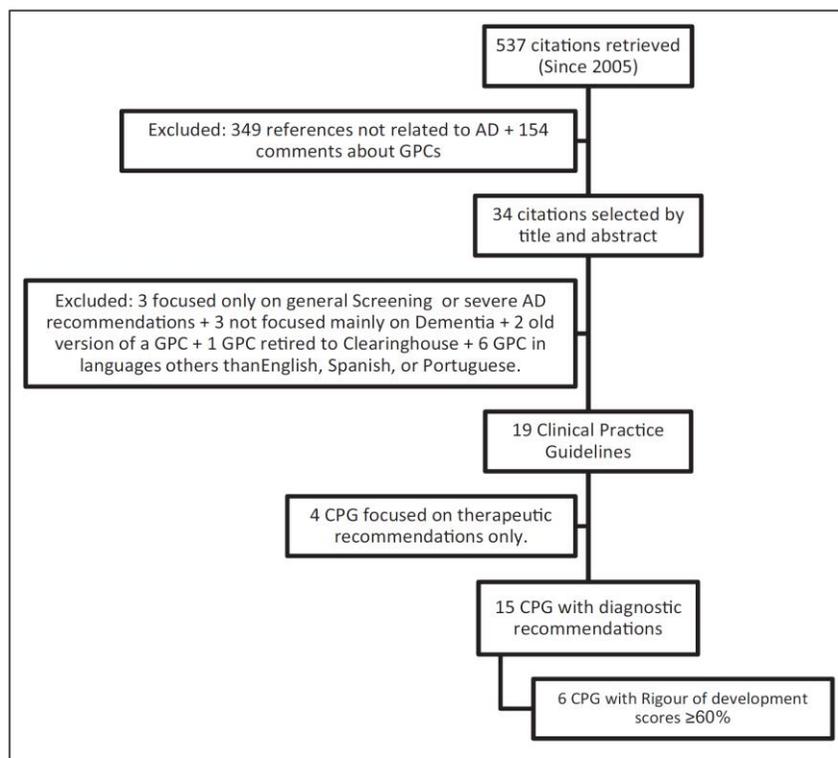


Figure 1. Flowchart of the systematic search.

1 CPG used a tool specifically designed for diagnostic questions³⁰ (Table 2). One CPG presented a diagnostic classification system in the Methods section, but after concluding that there was not enough information to address the diagnostic topics, they based their recommendations on an expert consensus and presented them without graduation of strength.¹⁸

All the 15 included guidelines had information on at least 1 of the 3 critical aspects of AD dementia diagnosis included in this review. In all, 14 (93%) provided recommendations related to the use of brief cognitive tests in the diagnosis of patients with suspected AD, 9 (60%) presented recommendations on the use of neuropsychological tests, and 4 (26%) presented explicit recommendations on use of β -amyloid accumulation/neuronal injury biomarkers.

Brief Cognitive Test Recommendations

All 6 (100%) high-quality guidelines provided recommendations on the use of brief cognitive tests in the diagnosis of patients with suspected AD dementia. They emphasized the importance of conducting a formal cognitive assessment using objective and valid instruments, although 3 (50%) of them did not recommend any test in particular (Table 3). A total of 19 brief instruments were recommended in the guidelines

retrieved, the Mini-Mental State Examination (MMSE) being the most frequently recommended. Of the 6 high-quality guidelines, 5 (83%) considered information on national validation and standardization of recommended tests, and 2 (33%) of them, recommended the use of MMSE versions validated in the country where the CPG had been developed.^{34,36}

Observational designs, especially cohorts and cross-sectional studies, were the most frequently included study designs. We found disagreements between the quality assessment of evidence and the confidence of recommendation among the high-quality guidelines. The former varied from very low to high depending on the system used. The strength of recommendation ranged from A to C (Table 3).

Neuropsychological Assessment Recommendations

All the 6 (100%) high-quality guidelines presented recommendations related to the use of neuropsychological tests. They recommended neuropsychological testing only when the diagnosis is unclear, when there is uncertainty about the stage of the disease, or when a differential diagnosis needs to be ruled out. Again, observational studies were the most frequently included type of design included. The quality of the evidence varied from very low to high and the strength of recommendations

Table 1. The AGREE II Domain-Standardized Scores for CPG on Alzheimer's Dementia Diagnostic Recommendations.

Guideline	Organization	Publication Year	Country and Language	Scope and Purpose	Stakeholder Involvement	Rigor of Development	Clarity of Presentation	Applicability	Editorial Independence
Diagnosis of Alzheimer's disease in Brazil ^{16,17}	Scientific Department of Cognitive Neurology and Aging of the Brazilian Academy of Neurology	2005	Brazil, Portuguese	64	26	57	88	18	15
Dementia: supporting people with dementia and their carers in health and social care ⁸	National Institute for Health and Care Institute for Excellence	2006	United Kingdom, English	92	88	83	86	65	77
Management of patients with dementia: A national clinical guideline ¹⁹	Scottish Intercollegiate Guidelines Network	2006	United Kingdom, English	69	64	92	93	93	58
Recommendations for best practices in the treatment of Alzheimer's disease in managed care ³¹	The Alzheimer's Drug Discovery Foundation	2006	United States, English	74	64	49	86	29	40
Dementia ²¹	Ministry of Health, Singapore	2007	Singapore, English	86	69	49	90	52	4
Cognitive impairment in the elderly: recognition, diagnosis, and management ²⁰	British Columbia Medical Association	2007	Canada, English	68	17	16	72	29	10
Diagnosis and treatment of dementia ²³⁻²⁹	Third Canadian Consensus Conference on the Diagnosis and Treatment of Dementia	2008	Canada, English	68	57	74	93	24	92
Guideline for Alzheimer's disease management ²²	California Workgroup on Guidelines for Alzheimer's Disease Management	2008	United States, English	76	53	38	75	22	19
Management of dementia ³⁶	Ministry of Health, Malaysia	2009	Malaysia, English	99	71	81	93	34	94
Dementia: diagnosis and treatment guideline ³³	Group Health Cooperative	2009	United States, English	26	22	18	49	11	8
Clinical practice guideline. Diagnosis and treatment of dementia in elderly in primary care ³⁵	Instituto Mexicano del Seguro Social	2009	México, Spanish	92	53	51	54	21	48
Clinical practice guideline about comprehensive care for people with Alzheimer's disease and other dementias ³⁴	Agència d'Informació, Avaluació i Qualitat en Salut de Catalunya	2010	Spain, Spanish	100	88	80	92	38	92
EFNS guidelines for the diagnosis and management of Alzheimer's disease ³⁰	European Federation of Neurological Societies	2010	International (Europe), English	65	47	70	88	28	83
Dementia: Diagnosis and Treatment ⁷	Regional Health Council—Tuscany	2011	Italy, English	75	63	27	79	3	15
Alzheimer's disease. Clinical practice guideline ³²	Argentine Neurological Society Research Group on Behavioral Neurology and Cognitive Neurosciences	2011	Argentina, Spanish	76	49	44	63	20	58
Mean scores				75.4	55.3	55.3	80.0	32.4	47.5
Range				26-100	17-88	16-92	49-93	3-93	4-94

Abbreviations: CPG, clinical practice guideline; EFNS, European Federation of Neurological Societies.

Table 2. Quality of Evidence and Strength of Recommendation Systems Used.

	Quality of the Evidence	Strength of Recommendation	CPG
SIGN ³⁸	<p>I⁺⁺ = High-quality meta-analyses, systematic reviews of randomized controlled trials (RCTs), or RCTs with a very low risk of bias.</p> <p>I⁺ = Well-conducted meta-analyses, systematic reviews of RCTs, or RCTs with a low risk of bias.</p> <p>2⁺⁺ = High-quality systematic reviews of case control or cohort studies; high-quality case-control or cohort studies with a very low risk of confounding or bias and a high probability that the relationship is causal.</p> <p>3 = Nonanalytic studies, eg, case reports, case series.</p> <p>4 = Expert opinion.</p>	<p>Level A = At least 1 meta-analysis, systematic review of RCTs, or RCT rated as I⁺⁺ and directly applicable to the target population; or a body of evidence consisting principally of studies rated as I⁺, directly applicable to the target population, and demonstrating overall consistency of results.</p> <p>Level B = A body of evidence including studies rated as 2⁺⁺, directly applicable to the target population, and demonstrating overall consistency of results; or Extrapolated evidence from studies rated as I⁺⁺ or I⁺.</p> <p>Level C = A body of evidence including studies rated as 2⁺, directly applicable to the target population and demonstrating overall consistency of results; or Extrapolated evidence from studies rated as 2⁺⁺.</p> <p>Level D = Evidence level 3 or 4; or extrapolated evidence from studies rated as 2⁺.</p>	19,34
Canadian Task Force on Preventive Health Care	<p>1 = Evidence obtained from at least 1 proper RCT.</p> <p>2 = Evidence obtained from well-designed controlled trials without randomization, or evidence obtained from well-designed cohort or case-control analytic studies preferably from more than 1 center or research group, or evidence obtained from comparisons between times or places with or without the intervention. Dramatic results in uncontrolled experiments are included in this category.</p> <p>3 = Opinions of respected authorities based on clinical experience, descriptive studies, or reports of expert committees.</p>	<p>Level A. There is good evidence to support this maneuver.</p> <p>Level B. There is fair evidence to support this maneuver.</p> <p>Level C. There is insufficient evidence to recommend for or against this maneuver, but recommendations may be made on other grounds.</p> <p>Level D. There is a fair evidence to recommend against this procedure.</p> <p>Level E. There is good evidence to recommend against this procedure.</p>	25
EFNS Classification	<p>I = A prospective study in a broad spectrum of persons with the suspected condition, using a "Gold Standard" for case definition, where the test is applied in a blinded evaluation, and enabling the assessment of appropriate tests of diagnostic accuracy.</p> <p>II = A prospective study of a narrow spectrum of persons with the suspected condition, or a well-designed retrospective study of a broad spectrum of persons with an established condition (by Gold Standard) compared to a broad spectrum of controls, where test is applied in a blinded evaluation, and enabling the assessment of appropriate tests of diagnostic accuracy.</p> <p>III = Evidence provided by a retrospective study where either persons with the established condition or controls are of a narrow spectrum, and where test is applied in a blinded evaluation.</p> <p>Class IV = Any design where test is not applied in blinded evaluation or evidence provided by expert opinion alone or in descriptive case series (without controls).</p>	<p>Level A rating (established as useful/predictive or not useful/predictive).</p> <p>Level B rating (established as probably useful/predictive or not useful/predictive).</p> <p>Level C rating (established as possibly useful/predictive or not useful/predictive) requires at least 2 convincing Class III studies.</p>	30
Canadian Task Force on Preventive Health Care and SIGN Modified	<p>I = Evidence obtained from at least 1 properly designed RCT.</p> <p>II-1 = Evidence obtained from well-designed controlled trials without randomization.</p> <p>II-2 = Evidence obtained from well-designed cohort or case-control analytic studies, preferably from more than 1 center or research group.</p> <p>II-3 = Evidence obtained from multiple time series with or without the intervention. Dramatic results in uncontrolled experiments could also be regarded as this type of evidence.</p> <p>III = Opinions of respected authorities, based on clinical experience; descriptive studies and case reports; or reports of expert committees.</p>	<p>Level A = At least 1 meta-analysis, systematic review, or RCT, or evidence rated as good and directly applicable to the target population</p> <p>Level B = Evidence from well-conducted clinical trials, directly applicable to the target population, and demonstrating overall consistency of results; or evidence extrapolated from meta-analysis, systematic review, or RCT</p> <p>Level C = Evidence from expert committee reports, or opinions and/or clinical experiences of respected authorities; indicates absence of directly applicable clinical studies of good quality.</p>	36

Abbreviations: EFNS, European Federation of Neurological Societies; SIGN, Scottish Intercollegiate Guidelines Network; RCT, randomized controlled trials.

Table 3. Recommendations for Brief Cognitive Tests in AD Dementia.

	N	Level of Evidence	Grade of Recommendation
MMSE and/or other tests	3	2 ^{++19,34} and III ³⁶	B ¹⁹ , C, ³⁶ and A ³⁴
Generic recommendation about brief cognitive test	3	2, ²⁵ I, ³⁰ and expert consensus ¹⁸	B, ²⁵ A, ³⁰ and no strength ¹⁸

Abbreviations: AD, Alzheimer's disease; MMSE, Mini-Mental State Examination.

Table 4. Neuropsychological Assessment Recommendations in AD Dementia Guidelines.

	Level of Evidence	Grade of Recommendation
"Fomal neuropsychological testing should form part of the assessment in cases of mild or questionable dementia" ¹⁸	Expert consensus	No strength
"Neuropsychological testing should be used in the diagnosis of dementia, especially in patients where dementia is not clinically obvious" ¹⁹	I ⁺⁺ , 2 ⁺⁺	B
"The diagnosis and differential diagnosis of dementia is currently a clinically integrative one. Neuropsychological testing alone cannot be used for this purpose and should be used selectively in clinical settings. Neuropsychological testing may aid in: addressing the distinction between normal aging, mild cognitive impairment or cognitive impairment without dementia, and early dementia; addressing the risk of progression from mild cognitive impairment or cognitive impairment without dementia to dementia or Alzheimer dementia; determining the differential diagnosis of dementia and other syndromes of cognitive impairment." ²⁵	2	B
"When the diagnosis of dementia is inconclusive, then neuropsychological tests will be required" ³⁶	III	C
"We recommend performed a detailed neuropsychological evaluation by specific tests when there are discrepancies between the clinical impression and screening tests, diagnostic concerns or when the complaints are of short duration or limited to a single cognitive domain" ³⁴	4	D
"Quantitative neuropsychological testing should be made in patients with questionable or very early Alzheimer Dementia" ³⁰ "The assessment of cognitive functions should include a general cognitive measure and more detailed testing of the main cognitive domains, and in particular an assessment of delay recall" ³⁰	III, I	B, A

Abbreviation: AD, Alzheimer's disease.

from A to D, similar to brief cognitive recommendations. None of the CPGs proposed a specific neuropsychological battery of instruments and only 1 guideline included a recommendation about the assessment of major cognitive domains (memory, executive functions, language, attention, and visuospatial skills; Table 4).³⁰

The AD Biomarkers Recommendations

Only 4 (66%) of the 6 high-quality CPGs presented explicit recommendations for the use of β -amyloid accumulation/neuronal injury biomarkers (Table 5). In all, 3 of these emphasized a lack of clinical studies on these biomarkers and presented recommendations against their use.^{19,34,36} The fourth high-quality CPG found FDG-PET and Tau levels to be useful,³⁰ although the direction of recommendation (for or against) remains unclear. Again, observational designs were the most frequently included studies. The quality of evidence varied widely, from low to high, depending on the system used. Similarly, the strength of recommendation ranged from B to C.

Discussion

This review shows that CPGs addressing diagnostic issues related to AD dementia published between 2005 and 2011 are very heterogeneous in terms of quality according to their AGREE-II scores. Domains such as applicability and editorial independence had very low scores and deserve special attention. Previous reviews of AD dementia CPGs highlight similar difficulties on these topics, even though such shortcomings could be easily improved.^{7,12} For example, if the applicability domain addresses organizational barriers, cost implications, and monitoring criteria related to guideline implementation,³⁹⁻⁴² it is important that guideline developers identify a priori factors related to transferability, such as values, costs, availability of resources, and influence of economic and intellectual interest on implementation. Regarding editorial independence issues,^{41,43,44} a complete and valid documentation of conflicts of interest for all participants in all stages of guideline development is required.⁴⁵

Issues concerning CPG rigor of development deserve closer attention. In our review, we noticed that 26% of the CPGs did

Table 5. Recommendations for the Use of Alzheimer's Dementia Biomarkers.

Level of evidence and strength of recommendation ^a	Biomarkers of A β Accumulation		Biomarkers of Neuronal Injury		
	Abnormal tracer retention on amyloid PET imaging	Low CSF A β 42	Elevated CSF Tau (total or phosphorylated)	Decreases FDG uptake on PET involving temporoparietal cortex	Atrophy on MRI, involving medial, basal and lateral temporal lobes and medial and lateral cortices.
Management of patients with dementia: A national clinical guideline ¹⁹		2 ⁺⁺ , B ^b	2 ⁺⁺ , B ^b		
Management of dementia ³⁶	III, C ^c	III, C ^b	III, C ^b	III, C ^b	III, C ^b
Clinical practice guidelines about comprehensive care for people with Alzheimer's disease and other dementias ³⁴		2 ⁺⁺ , B ^b	2 ⁺⁺ , B ^b		
EFNS guidelines for the diagnosis and management of Alzheimer's disease ³⁰		I, B ^c	I, B ^c	I, B ^c	I, B ^c

Abbreviations: CSF, cerebrospinal fluid; FDG, fluorodeoxyglucose; EFNS, European Federation of Neurological Societies; MRI, magnetic resonance imaging; PET, positron emission tomography.

^a Inside cells: level of evidence, grade of recommendation.

^b Recommendation against use of biomarker.

^c Recommendation unclear for use of biomarker.

not use a system to grade the quality of the evidence and the strength of recommendations. However, among those that did use a system, only 1 included diagnostic-specific elements in their framework (eg, appropriate reference standard or adequate spectrum of patients).⁴⁶ This figure is disturbing because CPG users need to know how much they can trust the evaluated information and whether adherence to the recommendations will yield more benefits than harm.⁴⁷ In our opinion, the use of generic systems developed by therapeutic studies would lead to inaccurate ratings of the quality of the evidence and would generate uncertainty about the recommendations. An example is the use of neuropsychological testing. Although recommendations were fairly homogeneous and specific in terms of content, their strength and quality of the evidence on which they were based varied from very low (experts opinion, D) to very high (1⁺⁺, A). This situation is worrying because although we selected only high-quality guidelines for these comparisons, it is not clear whether the evidence related to neuropsychological tests is of good quality or whether developers are confident about implementing the recommendations provided.

We found a similar situation regarding the use of brief cognitive tests, which added to other concerns about the clinical utility of these tools. In our review, we found a total of 19 tests recommended for initial assessment, but 3 of the high-quality guidelines only prescribed a cognitive assessment, although without identifying any instrument to achieve that goal. Once again, we encountered difficulties in assessing the quality of the diagnostic evidence and the strength of recommendations, but this variability also reflects doubts about the advantages and limitations of the tests.²⁵ We think that valid and global figures about accuracy of these tools, especially those related to subgroups such as patients with mild cognitive impairment

or the general population, are imperative. This information could help us determine which test(s) would be more useful for the initial assessment of patients with suspected AD dementia.

We found that only a few (26%) CPGs addressed the use of biomarkers. Three of them did not recommend their use in daily clinical practice and only one suggested that they could be potentially useful.³⁰ Based on the currently available evidence, the uncertainty with respect to the applicability and benefit of these biomarkers in clinical settings does not support their inclusion in CPGs. The CPG users, who still wish to adopt these diagnostic tools in daily practice, could face difficulties related to local standardization, development of cutoff points for different stages of AD, and availability of trained staff (eg, radiologists in measurement of MRI structures). Once again, we noticed that formal evaluation of clinical accuracy of these biomarkers is necessary.⁴⁸

A previous review by Beck et al also found considerable differences among diagnostic recommendations in CPGs on dementia published until 1999.¹¹ They attributed these differences to the methods used to formulate recommendations and to the need for an evidence-based approach for future guideline development. We found that disagreement persisted even in rigorously developed evidence-based CPGs, causing significant consequences in terms of interpretation, comparison, and generalizability of the recommendations.¹² In our opinion, adopting a system that took into account the nature of the evidence evaluated would be helpful. The framework outlined by Grading of Recommendations Assessment, Development and Evaluation approach has been developed to provide a common, rigorous, and comprehensive framework with a specific consideration of diagnostic issues that could be helpful in solving these difficulties.⁴⁹

Our review has several strengths and limitations. We conducted an exhaustive and systematic search strategy of the available literature that was not limited to indexed databases, given that it included references that had not been used in other reviews on AD dementia. However, we had to exclude 6 CPGs in languages such as German and Chinese as well as 4 references with only therapeutic recommendations, which could have facilitated a more complete evaluation of AD dementia guidelines. Similarly, we only included guidelines published between 2005 and 2011. Although we selected this time span in order to be more certain about the appropriateness of the guidelines, CPG developers may have had different interpretations of the same evidence.

We used a valid and widely disseminated instrument (AGREE-II) to assess guideline quality as well as a reliable process based on independent assessment by 4 appraisers. Nevertheless, the use of the AGREE instrument has potential drawbacks related to a distinctive focus on the development process of the CPG instead of on the quality of individual recommendations within the guideline.⁵⁰ In addition, we only used 1 domain of the AGREE-II tool to select the guidelines for individual analyses, whereas other reviews recommend the assessment of 2 domains, namely rigor of development and editorial independence. In our review, we aimed to assess the methodological issues related to the development of CPGs in order to highlight the issues associated with using generic grading systems, but we also noticed that all selected guidelines obtained adequate scores on the editorial independence domain ($\geq 50\%$).

As a conclusion, it is necessary that clinicians have access to high-quality and up-to-date recommendations from the best available evidence for the adequate management of patients with AD dementia. The quality of CPGs on the diagnosis of AD dementia can be improved not only through the inclusion of ethical aspects, the participation of stakeholders, and a clear strategy for their implementation, but also through a thorough evaluation of the quality of the available evidence using a unified framework to present recommendations.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Pablo Alonso-Coello is funded by a Miguel Servet investigator contract from the Instituto de Salud Carlos III. Other authors received no financial support for this research, authorship, and/or publication of this article.

References

1. Ferri CP, Prince M, Brayne C, et al. Global prevalence of dementia: a Delphi consensus study. *Lancet*. 2005;366(9503):2112-2117.
2. Wimo A, Winblad B, Jonsson L. The worldwide societal costs of dementia: Estimates for 2009. *Alzheimer's Dement*. 2010;6:98-103.
3. McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM. Clinical diagnosis of Alzheimer's disease: report 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(7):939-944.
4. McKhann GM, Knopman DS, Chertkow H, et al. The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's Dement*. 2011;7:263-269.
5. Dubois B, Feldman HH, Jacova C, et al. Revising the definition of Alzheimer's disease: a new lexicon. *Lancet Neurol*. 2010;9:1118-1127.
6. Woolf SH, Grol R, Hutchinson A, Eccles M, Grimshaw J. Clinical guidelines: potential benefits, limitations, and harms of clinical guidelines. *BMJ*. 1999;318(7182):527-530.
7. Alonso-Coello P, Irfan A, Sola I, et al. The quality of clinical practice guidelines over the last two decades: a systematic review of guideline appraisal studies. *Qual Saf Health Care*. 2010;19(6):e58.
8. Brouwers MC, Kho ME, Browman GP, et al. AGREE II: advancing guideline development, reporting and evaluation in health care. *J Clin Epidemiol*. 2010;63(12):1308-1311.
9. Ferket BS, Colkesen EB, Visser JJ, et al. Systematic review of guidelines on cardiovascular risk assessment: which recommendations should clinicians follow for a cardiovascular health check? *Arch Intern Med*. 2010;170(1):27-40.
10. Bennett WL, Odelola OA, Wilson LM, et al. Evaluation of guideline recommendations on oral medications for type 2 diabetes mellitus: a systematic review. *Ann Intern Med*. 2012;156(1 pt 1):27-36.
11. Beck C, Cody M, Souder E, Zhang M, Small GW. Dementia diagnostic guidelines: methodologies, results, and implementation costs. *J Am Geriatr Soc*. 2000;48(10):1195-1203.
12. Azemai M, Petrovic M, Elseviers MM, Bourgeois J, Van Bortel LM, Vander Stichele RH. Systematic appraisal of dementia guidelines for the management of behavioural and psychological symptoms. *Ageing Res Rev*. 2011;11:78-86.
13. Cummings JL, Frank JC, Cherry D, et al. Guidelines for managing Alzheimer's disease: Part II. Treatment. *Am Fam Physician*. 2002;65(12):2525-2534.
14. Cummings JL, Frank JC, Cherry D, et al. Guidelines for managing Alzheimer's disease: part I. Assessment. *Am Fam Physician*. 2002;65:2263-2272.
15. Watine J, Friedberg B, Nagy E, et al. Conflict between guideline methodologic quality and recommendation validity: a potential problem for practitioners. *Clin Chem*. 2006;52(1):65-72.
16. Nitrini R, Caramelli P, Bottino CM, Damasceno BP, Brucki SM, Anghinah R. [Diagnosis of Alzheimer's disease in Brazil: diagnostic criteria and auxiliary tests. Recommendations of the Scientific Department of Cognitive Neurology and Aging of the Brazilian Academy of Neurology]. *Arq Neuropsiquiatr*. 2005; 63(3A):713-719.
17. Nitrini R, Caramelli P, Bottino CM, Damasceno BP, Brucki SM, Anghinah R. [Diagnosis of Alzheimer's disease in Brazil: cognitive and functional evaluation. Recommendations of the Scientific Department of Cognitive Neurology and Aging of the Brazilian Academy of Neurology]. *Arq Neuropsiquiatr*. 2005; 63(3A):720-727.

18. National Collaborating Centre for Mental Health, National Institute for Health and Clinical Excellence (NICE). Dementia: supporting people with dementia and their carers in health and social care. London, England: National Institute for Health and Clinical Excellence (NICE); 2006:417.
19. Scottish Intercollegiate Guidelines Network. *Management of patients with dementia. A national clinical guideline*. Edinburgh, Scotland: Scottish Intercollegiate Guidelines Network (SIGN); 2006:53.
20. British Columbia Medical Association. Cognitive Impairment in the Elderly-Recognition, Diagnosis and Management. Clinical Practice Guidelines and Protocols in British Columbia: 2007 [cited 2011 July 01]. <http://www.bcguidelines.ca/pdf/cognitive.pdf>.
21. Singapore Ministry of Health. *Dementia*. Singapore: Singapore Ministry of Health; 2007:80.
22. California Workgroup on Guidelines for Alzheimer's Disease Management. Guideline for Alzheimer's disease management. Chicago, IL: Alzheimer's Association; 2008:61.
23. Chertkow H. Diagnosis and treatment of dementia: introduction. Introducing a series based on the Third Canadian Consensus Conference on the Diagnosis and Treatment of Dementia. *CMAJ*. 2008;178(3):316-321.
24. Chertkow H, Massoud F, Nasreddine Z, et al. Diagnosis and treatment of dementia: 3. Mild cognitive impairment and cognitive impairment without dementia. *CMAJ*. 2008;178(10):1273-85.
25. Feldman HH, Jacova C, Robillard A, et al. Diagnosis and treatment of dementia: 2. Diagnosis. *CMAJ*. 2008;178(7):825-836.
26. Herrmann N, Gauthier S. Diagnosis and treatment of dementia: 6. Management of severe Alzheimer disease. *CMAJ*. 2008;179(12):1279-1287.
27. Hogan DB, Bailey P, Black S, et al. Diagnosis and treatment of dementia: 4. Approach to management of mild to moderate dementia. *CMAJ*. 2008;179(8):787-793.
28. Hogan DB, Bailey P, Black S, et al. Diagnosis and treatment of dementia: 5. Nonpharmacologic and pharmacologic therapy for mild to moderate dementia. *CMAJ*. 2008;179:1019-1026.
29. Patterson C, Feightner JW, Garcia A, Hsiung GY, MacKnight C, Sadvnick AD. Diagnosis and treatment of dementia: 1. Risk assessment and primary prevention of Alzheimer disease. *CMAJ*. 2008;178(5):548-556.
30. Hort J, O'Brien JT, Gainotti G, et al. EFNS guidelines for the diagnosis and management of Alzheimer's disease. *Eur J Neurol*. 2010;17(10):1236-1248.
31. Fillit HM, Doody RS, Binaso K, et al. Recommendations for best practices in the treatment of Alzheimer's disease in managed care. *Am J Geriatr Pharmacother*. 2006;4(Suppl A):S9-S24;quiz S5-S8.
32. Allegri RF, Arizaga RIL, Bavec CV, et al. Enfermedad de Alzheimer. Guía de práctica clínica. *Neurol Argent*. 2011;3(2):120-137.
33. Group Health Cooperative. *Dementia*. Diagnosis and treatment Guideline. Seattle, Wash; 2009:13.
34. Grupo de trabajo de la Guía de práctica clínica sobre la atención integral a las personas con enfermedad de Alzheimer y otras demencias. Guía de práctica clínica sobre la atención integral a las personas con enfermedad de Alzheimer y otras demencias. Agència d'Informació, Avaluació i Qualitat en Salut. Servei Català de la Salut. Pla Director Sociosanitario. Departament de Salut. Generalitat de Catalunya; 2011:508.
35. Instituto Mexicano del Seguro Social. *Guía de Práctica Clínica: Diagnóstico y Tratamiento de la Demencia en el Adulto Mayor en el Primer Nivel de Atención*. México: Instituto Mexicano del Seguro Social; 2010:45.
36. Ministry of Health Malaysia. *Management of Dementia*. 2nd ed. Malaysia: Ministry of Health Malaysia; 2009:144.
37. Regional Health Council. *Dementia. Diagnosis and treatment*. Milan, Italy: Regione Toscana, Consiglio Sanitario Regionale; 2011:38.
38. Scottish Intercollegiate Guidelines Network. SIGN 50: A guideline developer's handbook. Edinburgh, England: 2008 [cited 2012 Jan 01]. <http://www.sign.ac.uk/guidelines/fulltext/50/index.html>.
39. Schunemann HJ, Fretheim A, Oxman AD. Improving the use of research evidence in guideline development: 13. Applicability, transferability and adaptation. *Health Res Policy Syst*. 2006;4:25.
40. Fretheim A, Schunemann HJ, Oxman AD. Improving the use of research evidence in guideline development: 15. Disseminating and implementing guidelines. *Health Res Policy Syst*. 2006;4:27.
41. Lexchin J, Bero LA, Djulbegovic B, Clark O. Pharmaceutical industry sponsorship and research outcome and quality: systematic review. *BMJ*. 2003;326:1167-1170.
42. Francke AL, Smit MC, de Veer AJ, Mistiaen P. Factors influencing the implementation of clinical guidelines for health care professionals: a systematic meta-review. *BMC Med Inform Decis Mak*. 2008;8:38.
43. Boyd EA, Bero LA. Improving the use of research evidence in guideline development: 4. Managing conflicts of interests. *Health Res Policy Syst*. 2006;4:16.
44. Detsky AS. Sources of bias for authors of clinical practice guidelines. *CMAJ*. 2006;175(9):1033,1055.
45. Guyatt G, Akl EA, Hirsh J, et al. The vexing problem of guidelines and conflict of interest: a potential solution. *Ann Intern Med*. 2010;152(11):738-41.
46. Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med*. 2011;155(8):529-536.
47. Schunemann HJ, Fretheim A, Oxman AD. Improving the use of research evidence in guideline development: 9. Grading evidence and recommendations. *Health Res Policy Syst*. 2006;4:21.
48. McShane R, Noel-Storr A, Ritchie C, Flicker L. The quality and extent of evidence for biomarkers: a Cochrane systematic review. *Alzheimer's & dementia: the journal of the Alzheimer's Association*. 2011;7(4):S100-S1.
49. Schunemann HJ, Oxman AD, Brozek J, et al. GRADE: assessing the quality of evidence for diagnostic recommendations. *Evid Based Med*. 2008;13(6):162-163.
50. Brouwers MC, Kho ME, Browman GP, et al. The Global Rating Scale complements the AGREE II in advancing the quality of practice guidelines. *J Clin Epidemiol*. 2012;65(5):526-534.

Diagnostic tools for alzheimer's disease dementia and other dementias: an overview of diagnostic test accuracy (DTA) systematic reviews.

Ingrid Arevalo-Rodriguez, Omar Segura, Ivan Solà, Xavier Bonfill, Erick Sanchez and Pablo Alonso-Coello.

BMC Neurology 2014, 14:183

Factor de impacto: 2.49.

RESUMEN DE RESULTADOS

Se identificaron un total de 549 referencias después de la eliminación de duplicados, seleccionándose para lectura en texto completo un total de 76 documentos. 52 de dichos documentos fueron excluidos debidos a que presentaban una revisión narrativa referente a la demencia en general o no contenían suficientes detalles para la valoración completa de su calidad (resúmenes de congresos). Al final del proceso, se seleccionaron 24 revisiones sistemáticas de exactitud diagnóstica (100-123), las cuales contaban con un tamaño muestral mediano de 2.190 pacientes (Desde 160 a 26.019).

Diez de estas revisiones (41.6%) se enfocaron en el diagnóstico de DCL, mientras nueve revisiones (37.5%) abordaron el diagnóstico de demencia por EA, con las restantes ocho revisiones abordando la demencia en general. Ocho de estas revisiones (33%) abordaron más de un tipo de demencia. Siete referencias (29%) incluyeron un tamaño muestral inferior a 1000 pacientes, mientras que un número similar de revisiones no reportó el número total de pacientes derivado de los estudios primarios analizados. Nueve revisiones (37.5%) incluyeron menos de 10 estudios de exactitud diagnóstica.

Las revisiones seleccionadas incluyeron en su mayoría estudios de corte transversal o cohortes, con una mediana de 19.5 estudios primarios analizados (desde 2 hasta 233 estudios de diagnóstico analizados). La prueba diagnóstica más evaluada fueron las pruebas cognitivas breves (nueve revisiones), seguida por el PET/SPECT y los niveles de Tau en LCR (Seis revisiones cada una). Los autores de las revisiones emplearon diferentes estándares de referencia para validar el

diagnóstico de demencia, siendo los criterios NINCS-ARDRA y los del DSM-IV los más frecuentemente reportados (once y nueve revisiones, respectivamente); cuatro de las revisiones identificadas no indican explícitamente el estándar de referencia empleado.

Respecto a la evaluación de la calidad metodológica de los estudios primarios, solo 10 revisiones (41.6%) reportan dicha valoración, siendo el instrumento QUADAS-I la herramienta más usada para la valoración del riesgo de sesgo de dichos estudios (6 estudios, 60%). Los sesgos más frecuentemente reportados fueron el sesgo en el espectro de pacientes y el sesgo de incorporación. Se encontró que cuatro revisiones (16%) emplearon la herramienta STARD para calificar la calidad de los estudios analizados. No se encontró dentro de las revisiones reporte de resultados no-conclusivos, eventos adversos relacionados con las pruebas evaluadas o uso de recursos. Asimismo, solo once revisiones (45.8%) reportaron de manera explícita las fuentes de financiación de la revisión diagnóstica, siendo en su mayoría de origen gubernamental.

Respecto a la evaluación de la calidad del reporte por medio de la lista de chequeo PRISMA, todas las revisiones diagnósticas identificadas presentaron una justificación para el desarrollo de la revisión (100%), mientras 20 revisiones (83.3%) se identificaron en su título como “Revisiones sistemáticas” (Ítem 1-PRISMA). 91.6% de las revisiones reportaron gráficamente el número de estudios tamizados, evaluados por criterios de inclusión/ exclusión y aquellos incluidos (Ítem 17-PRISMA), mientras 87.5% presentó las características de los estudios incluidos con sus respectivas referencias (Ítem 18-PRISMA). Por el contrario, solo una revisión reportó la existencia de un protocolo previo a su desarrollo (Ítem 5-PRISMA), cuatro reportaron análisis adicionales (Ítem 23-PRISMA) y cinco presentaron una valoración del sesgo de publicación o reporte selectivo entre los estudios (Ítem 22-PRISMA). Solo un estudio reportó todos los ítems metodológicos sugeridos por la herramienta PRISMA (Ítems 5 a 16).

Finalmente, con respecto a la calidad de la revisión sistemática en términos de la herramienta AMSTAR (acorde con las definiciones operativas previamente presentadas en la sección de Metodología), 87.5% de las revisiones no proporcionaron una lista de estudios incluidos/excluidos, y 66.6% de la selección de estudios y la extracción de datos no fue realizada por duplicado. Seis de las revisiones identificadas emplearon análisis estadísticos desarrollados para revisiones de

efectividad (25%), mientras en siete revisiones no fue posible determinar el análisis estadístico empleado para la estimación conjunta de las características operativas de la prueba analizada. Asimismo, en cinco revisiones (20.8%) se evaluó el sesgo de publicación por medio de gráficas de Funnel Plot y de pruebas estadísticas como el test de Egger.

RESEARCH ARTICLE

Open Access

Diagnostic tools for alzheimer’s disease dementia and other dementias: an overview of diagnostic test accuracy (DTA) systematic reviews

Ingrid Arevalo-Rodriguez^{1,2*}, Omar Segura¹, Ivan Solà³, Xavier Bonfill⁴, Erick Sanchez⁵ and Pablo Alonso-Coello⁴

Abstract

Background: Dementia includes a group of neurodegenerative disorders characterized by progressive loss of cognitive function and a decrease in the ability to perform activities of daily living. Systematic reviews of diagnostic test accuracy (DTA) focus on how well the index test detects patients with the disease in terms of figures such as sensitivity and specificity. Although DTA reviews about dementia are essential, at present there is no information about their quantity and quality.

Methods: We searched for DTA reviews in MEDLINE (1966–2013), EMBASE (1980–2013), *The Cochrane Library* (from its inception until December 2013) and the Database of Abstracts of Reviews of Effects (DARE). Two reviewers independently assessed the methodological quality of the reviews using the AMSTAR measurement tool, and the quality of the reporting using the PRISMA checklist. We describe the main characteristics of these reviews, including basic characteristics, type of dementia, and diagnostic test evaluated, and we summarize the AMSTAR and PRISMA scores.

Results: We selected 24 DTA systematic reviews. Only 10 reviews (41.6%), assessed the bias of included studies and few (33%) used this information to report the review results or to develop their conclusions. Only one review (4%) reported all methodological items suggested by the PRISMA tool. Assessing methodology quality by means of the AMSTAR tool, we found that six DTA reviews (25%) pooled primary data with the aid of methods that are used for intervention reviews, such as Mantel-Haenszel and separate random-effects models (25%), while five reviews (20.8%) assessed publication bias by means of funnel plots and/or Egger’s Test.

Conclusions: Our assessment of these DTA reviews reveals that their quality, both in terms of methodology and reporting, is far from optimal. Assessing the quality of diagnostic evidence is fundamental to determining the validity of the operating characteristics of the index test and its usefulness in specific settings. The development of high quality DTA systematic reviews about dementia continues to be a challenge.

Keywords: Diagnosis, Dementia, Alzheimer’s disease dementia, Systematic review, PRISMA checklist, AMSTAR tool

Background

Population ageing is generating a considerable increase in chronic and neurodegenerative diseases, as well as severe consequences for global public health [1,2]. Dementia includes a group of neurodegenerative disorders characterized by progressive loss of cognitive function as well as

the ability to perform activities of daily living, sometimes accompanied by neuropsychiatric symptoms [3]. Criteria for dementia diagnosis include a deficit in one or more cognitive domains that is severe enough to impair functional activities, and is progressive over a period of at least six months and not attributable to any other brain disease [4,5]. The presence of cognitive impairment, a fundamental part of the dementia profile, could be detected through a combination of history, clinical examination, and objective cognitive assessment such as a brief mental assessment or comprehensive neuropsychological testing [6,7]. At present,

* Correspondence: iarevalo@fucsalud.edu.co

¹Division of Research, Fundación Universitaria de Ciencias de la Salud, Hospital de San José/Hospital Infantil de San José, Bogotá, DC, Colombia

²Pediatrics, Obstetrics & Gynecology and Preventive Medicine Department, Universitat Autònoma de Barcelona, Barcelona, Spain

Full list of author information is available at the end of the article



© 2014 Arevalo-Rodriguez et al.; licensee BioMed Central Ltd. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly credited. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated.

there is a trend towards incorporating biomarker tests into dementia diagnosis criteria, such as amyloid- β protein accumulation, neuronal injury, synaptic dysfunction, and neuronal degeneration [8-10].

Systematic reviews of diagnostic test accuracy (DTA) focus on how well an index test detects patients with the disease in terms of figures such as sensitivity and specificity. DTA reviews present summarized information to consumers (such as clinicians, stakeholders, guideline developers and patients) about which test should be used over another as the initial step in a diagnostic pathway or as an add-on element to confirm the presence of the target disease. Although the methodology for performing DTA reviews is constantly evolving, organizations such as the Cochrane Collaboration have published methodological guidance with basic requirements to develop these kinds of reviews [11].

Recently, we evaluated the quality of clinical practice guidelines for diagnosing dementia and found a wide variety in terms of quality of evidence as well as the strength of the recommendations provided [12]. Although DTA reviews are an essential part of any clinical guidelines, at present there is no information about the quantity and quality of dementia DTA reviews. This information could help clinicians and stakeholders provide adequate management and appropriate care for these patients, in line with the rise in dementia and its expected burden on health systems.

The objective of this study was to evaluate the quality (in terms of rigor in conduct and reporting) of DTA systematic reviews related to diagnostic tools for Alzheimer's disease dementia (ADD) and other dementias. These tools included brief cognitive tests, biomarkers, and neuropsychological assessment, and they were assessed by means of standardized tools, as well as by describing the tests evaluated and their main characteristics.

Methods

We produced a protocol for the review (available from the authors on request) detailing the proposed review methods. We searched in MEDLINE (1966–2013), EMBASE (1980–2013), *The Cochrane Library* (from its inception until December 2013) and the Database of Abstracts of Reviews of Effects (DARE), by means of a predesigned search strategy adapted to each database (Additional file 1), in order to identify diagnostic systematic reviews focused on the test accuracy of diagnostic tools for dementia, ADD or other dementias (e.g. vascular dementia, frontotemporal dementia, Lewy bodies, and Parkinson dementia). We checked the reference lists of the selected studies for additional references, and excluded congress abstracts and references with insufficient information.

Two reviewers independently assessed the eligibility of the results and extracted data from the selected studies.

In this overview we included systematic reviews of diagnostic studies that focused on the accuracy of tests for dementia. Only reviews that used a systematic approach, included adult patients aged over 50 suspected of having dementia, and estimated the accuracy of the assessed test (i.e. providing sensitivity and specificity figures) were considered. We used a predefined data extraction form to extract descriptive information including year of publication, type of studies included, and clinical reference standard, and whether a checklist was used to evaluate the methodological quality of primary studies (such as the Quality Assessment of Diagnostic Accuracy Studies, QUADAS [13,14]).

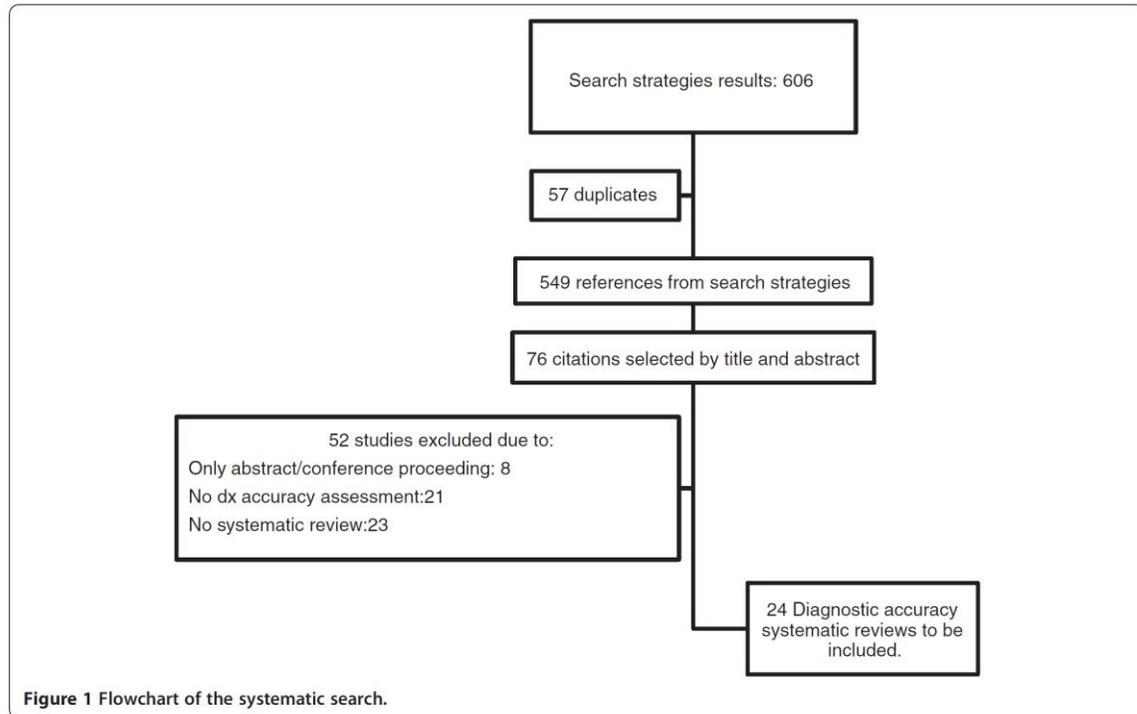
After this was done, two reviewers independently assessed the methodological quality of the selected reviews using the Assessment of Multiple Systematic Reviews (AMSTAR) measurement tool [15], tailored to the characteristics of DTA systematic reviews (Additional file 1). They also assessed the quality of the reporting using the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) checklist [16]. We resolved disagreements through discussion. In this article we describe the main characteristics of the selected reviews, including basic characteristics (e.g. reference standard used and diagnostic bias reported). We also describe the type of dementia and index test evaluated, as well as AMSTAR and PRISMA scores by item.

Results

We retrieved a total of 549 citations after excluding duplicates and initially selected a total of 76 references for full review. We excluded 52 articles because they did not provide diagnostic accuracy information, presented a narrative overview about dementia, or did not have enough information to evaluate their quality (i.e. congress abstracts) (Additional file 1). Finally, we selected 24 DTA systematic reviews [17-40], with a median sample size of 2,190 patients (range from 160 to 26,019) (Figure 1 and Additional file 1).

Ten reviews (41.6%) focused on mild cognitive impairment (MCI), an early stage of dementia, either for detection or conversion to full dementia, while nine reviews (37.5%) focused on ADD, and eight on dementia in general. Eight out of 24 DTA reviews (33%) included more than one subtype of dementia. Seven studies (29%) had less than 1,000 patients, and a similar number did not report the total number of patients derived from primary studies, while nine (37.5%) of the reviews included less than 10 studies (Additional file 1).

The reviews selected included mostly cross-sectional and cohort studies, with a median of 19.5 primary studies included (range from two to 233 primary studies). The index tests most frequently evaluated were cognitive tests (nine DTA reviews), followed by PET/SPECT and



serum levels of Total Tau and P-Tau (six DTA reviews each). Several reference standards were used to validate dementia diagnoses, with NINCS-ARDRA and DSM-IV being the most common (11 and nine reviews, respectively). Four DTA reviews did not indicate the reference standard used to evaluate the validity of dementia diagnoses. Table 1 shows the selected reviews by type of dementia and diagnostic tool evaluated.

Only 10 reviews (41.6%) assessed the methodological quality of primary studies, with QUADAS-I being the most commonly used tool for assessing risk of bias of primary studies (six studies, 60%). Patient spectrum and incorporation bias were the most frequent biases reported by review authors. Four reviews assessed the methodological quality of primary studies by means of the STARD tool (16%), which is intended to assess reporting quality. None of the DTA reviews reported results related to inconclusive results, adverse events, or the use of resources related to index tests in an explicit way. Eleven reviews (45.8%) reported the sources of funding or support to perform the DTA review with most of them being government sources.

With regards to the PRISMA checklist, all selected reviews (100%) described the rationale for the review (Item 3), and 20 (83.3%) identified themselves as systematic reviews (Item 1). Twenty-two reviews (91.6%) reported the number of studies screened, assessed for eligibility,

and included in the review, by means of a flow chart (Item 17) and 21 (87.5%) presented characteristics of studies and provided citations (Item 18). However, only one study (4.1%) reported a review protocol (Item 5), four (16.6%) reported results of additional analysis (Item 23), and five (20.8%) presented results of risk of bias assessment across studies as publication bias or selective reporting within studies (Item 22). Only one review reported all methodological items suggested by the PRISMA tool (Items 5 to 16) (Figure 2).

With respect to the quality of conduct in terms of the AMSTAR tool, 21 reviews (87.5%) did not provide a list of included/excluded studies and 16 (66.6%) did not report duplicate study selection/data extraction (Figure 3). All reviews reported the characteristics of included studies (100%). Six DTA reviews (25%) pooled primary data by means of methods that are used for intervention reviews such as Mantel-Haenszel and separate random-effects models, while in seven reviews (29%) it was not possible to determine which methods were used to combine the numerical results. Five reviews (20.8%) assessed publication bias by means of funnel plots and/or Egger's Test. Fourteen of these DTA reviews (58.3%) reported possible conflicts of interest. As mentioned above, only 10 reviews (41.6%) assessed the bias of included studies, and only eight (33%) used this information to report the review results or reach their conclusions.

Table 1 DTA systematic reviews about dementia by type of dementia and diagnostic tool evaluated

	MMSE	Other cognitive tests	PET/ SPECT	CSF Aβ ₄₂	P -Tau/T-Tau	FDG/PIB uptake on PET	MRI/ CT	Other diagnostic tools
ADD			Bloudek [19] Dougall [22] Ferrante [24]	Bloudek [19]	Bloudek [19] Mitchell [28]	Bloudek [19] Matchar [27] Patwardhan [34]	Bloudek [19]	
DLB			Yeo [40]		Van Harten [36]			Papathanasiou [33] Treglia [35] Yeo [40]
VaD			Dougall [22] Yeo [40]		Van Harten [36]		Beynon [18]	Yeo [40]
FTD			Dougall [22] Yeo [40]		Van Harten [36]			Yeo [40]
Dementia in general	Mitchell [29]	Appels [17] Carnero [20] Crawford [21] Mitchell [30] Mitchell [31]	Ferrante [24]			Matchar [27]		
MCI	Lischka [25] Lonie [26] Mitchell [29]	Ehreke [23] Lischka [25] Lonie [26]	Yuan [38]	Monge [32] van Rossum [37]	Mitchell [28] Monge [32] van Rossum [37]	Zhang [39]	Yuan [38]	

Abbreviations: Aβ₄₂ 42 aminoacid form of amyloid-β, ADD Alzheimer's Disease Dementia, CT Computed tomography, DLB Dementia with Lewy Bodies, FDG-PET PET using 2-Fluoro-deoxy D-glucose, FTD Fronto-Temporal Dementia, MCI Mild cognitive Impairment, MMSE Mini-Mental State Examination, MRI Magnetic Resonance Imaging, PIB-PET 11 C-Pittsburgh Compound B- positron emission tomography, Ptau Phosphorylated Tau, SPECT Single photon emission computed tomography, Ttau Total Tau, VaD Vascular Dementia.

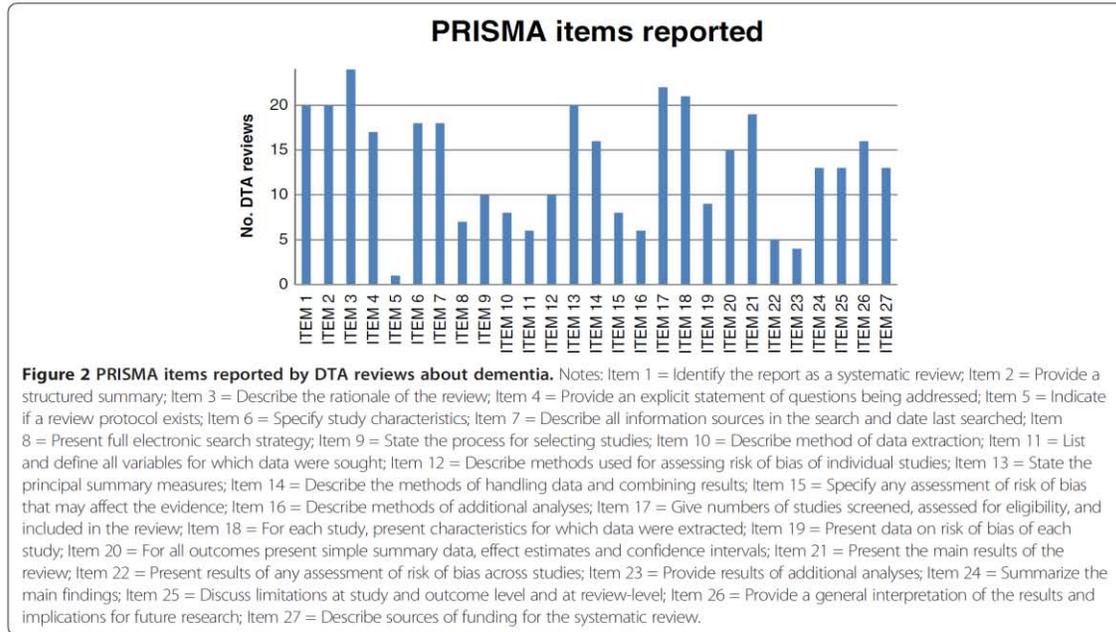
Discussion

Our review of DTA systematic reviews about dementia shows several areas for improvement. First, we had to exclude a significant number of reviews focused not on the accuracy of the test (i.e. sensitivity and specificity figures), but instead presenting information about the average differences between case and control groups. In these reviews, the authors gathered information about Phase I diagnostic studies, evaluating the differences (for example, in terms of difference of means) between a group of subjects with the disease and healthy controls [41]. These studies are essential for an adequate and full assessment of any diagnostic tool, but cannot show if the test distinguishes between those with and without the target condition. Authors of future reviews should be careful in appraising these studies due to the higher risk of bias (for instance, the wide use of cases and controls design) and their limitations in decision-making processes.

In relation to the basic characteristics of dementia DTA systematic reviews, we noticed that a significant number of reviews were focused on mild cognitive impairment (MCI). Identification of early forms of dementia has become an important topic because some interventions have been claimed to be effective in slowing or stopping the cognitive decline when they are administered in earlier stages of dementia, but these

findings are still being investigated [42-45]. Similarly, it is interesting that a significant number of reviews focused on cognitive tests, which are the first line of detection for cognitive impairment in dementia. At present it is unclear which cognitive test should be the instrument of choice for initial dementia screening in population-based, primary and secondary settings, due to rising criticism of the role of traditional tests such as Mini-Mental State examination (MMSE) [12].

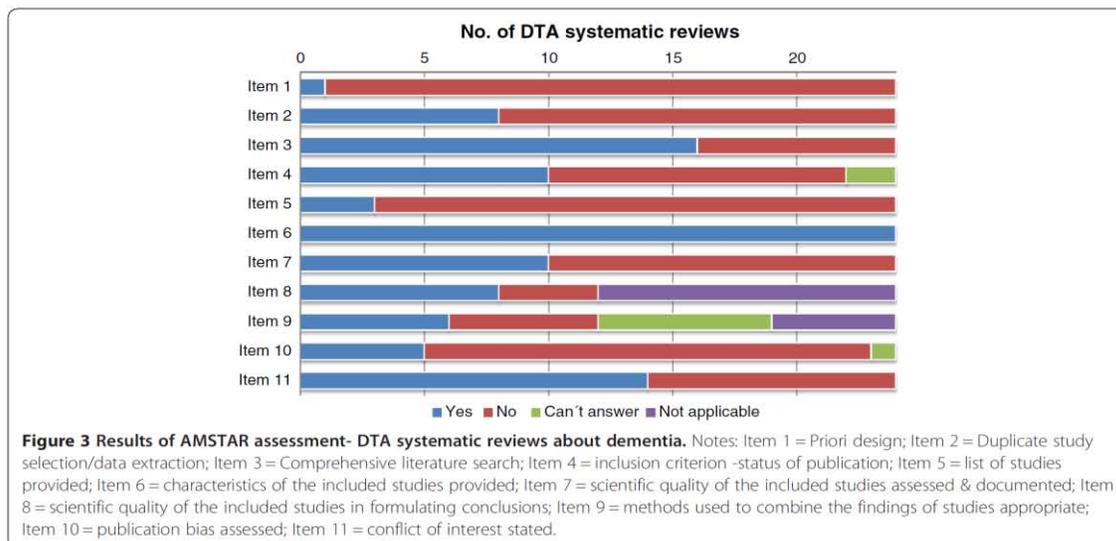
Our assessment of these DTA reviews reveals that their quality, in terms of both methodology and reporting, is far from optimal. We found that more than half of the included reviews did not provide a quality assessment of the primary studies, and therefore information of an unknown quality was gathered and even numerical pooled results were provided. Assessing the quality of diagnostic evidence is fundamental to determining the validity of the operating characteristics of the index test, and its usefulness in specific settings [46]. Four reviews (16%) did not report the reference standards they used to evaluate the accuracy of the different tests appraised, while others reported STARD scores as an evaluation of methodological quality. Only 13 of the 24 reviews (54%) described the limitations of the information gathered, and in only eight cases (33%) was the quality of this information considered in the conclusions. Likewise,



AMSTAR and PRISMA items in conjunction showed an almost complete absence of a priori protocols presenting pre-specified methodological plans. The importance of pre-specified protocols has been established in intervention reviews as well as in clinical trials of pharmacological interventions. Diagnostic tests for dementia, such as FDG-PET and Tau-AB₄₂, can be understood as medical

technologies that can be affected by conflicts of interest. The availability of protocols at the beginning of any study not only ensures rigor in development, but also avoids conducting unnecessary research [47].

In our study we also identified drawbacks in developing DTA systematic reviews related to the application of statistical methods generally used in intervention



systematic reviews. For example, when the methods used for pooling numerical information were assessed, we identified three reviews that used Der Simonian-Laird random effects models, instead of methods highly recommended in these cases such as bivariate models [48]. Some authors have asserted that the use of inadequate statistical techniques to deal with diagnostic information could lead to failures in managing the combined results of sensitivity and specificity [48-51]. Similarly, some reviews used the I^2 statistic to illustrate the heterogeneity between the analyzed studies. Heterogeneity is a common issue in accuracy reviews, due to factors such as threshold used, prevalence of the target condition in the sample selected, and settings of test evaluation [52], but at present, there are no defined standards for how diagnostic heterogeneity should be measured and managed in DTA reviews [53].

In the same way, we identified that evaluation of publication bias remains a problem in dementia DTA reviews. In our study, 18 reviews (75%) did not provide information about this bias, but it is unclear if the authors simply omitted this evaluation or if they decided not to assess this topic due to lack of suitable analysis methods. Three additional cases (12.5%) used funnel-plot figures or statistical tests (for instance, Egger's test). While these methods are highly useful in intervention systematic reviews, several research studies have shown that their use in the field of DTA reviews, usually by means of diagnostic odds ratios (DOR), can generate misleading results [54].

Our study has some limitations. One of these is related to the tools used to evaluate systematic reviews (such as PRISMA and AMSTAR), focused mainly on intervention reviews. In order to correctly use the AMSTAR tool we developed tailored definitions to adequately reflect the most accepted methodology of DTA systematic reviews. However, it is important to encourage discussion about how current tools and methodologies (for example, overview methodology) can be applied or adapted to developing DTA studies. A second difficulty that we found was the large number of diagnostic reviews reported only in abstract form, which had to be excluded because of the absence of information to allow for a full assessment of their elements. We believe that these "ongoing" studies reflect the growing interest in the diagnosis of dementia, as well as the need for comprehensive discussion about dementia diagnosis tools. Finally, our search strategy was specific and did not include terms related to MCI. Our findings related to this early stage of dementia might be incidental and not reflect all possible DTA reviews in this area.

Conclusions

Development of systematic reviews of diagnostic test accuracy for dementia remains a difficult task. However,

an increasing number of health professionals require information about the quality of diagnostic technologies due to their role in detecting, staging and monitoring. We believe that some recent initiatives might help improve methodology and reporting quality in DTA reviews on dementia [11,48,55]. In the near future, high quality DTA reviews could play an important role in helping clinicians, policy-makers and even patients to make informed decisions for the diagnosis of this prevalent disease.

Additional file

Additional file 1: This file includes the following information: 1. Search strategy used on MEDLINE. 2. Operational definitions of AMSTAR's items for DTA reviews about dementia. 3. DTA systematic reviews about dementia- Descriptive information. 4. DTA systematic reviews about dementia- AMSTAR items. 5. DTA systematic reviews about dementia- PRISMA items. 6. List of excluded studies.

Abbreviations

7MS: Seven-minute screen; ^{99m}Tc -HMPAO 99 m: Technetium-hexamethyl-propylenamine oxime; $\text{A}\beta_{1-42}$: 42 aminoacid form of amyloid- β ; ACE: Addenbrooke's cognitive examination; ADAS-CoG: Alzheimer's disease assessment scale- cognitive; ADD: Alzheimer's disease dementia; ADTC: State of California AD diagnostic and treatment centre criteria; BKSCAR: Brief Kingston standardized cognitive assessment- revised; BVRT: Benton's visual retention test; CAMCI: Chinese abbreviated mild cognitive impairment test; CAMCOG: Cambridge cognitive examination; CASI: Cognitive abilities screening instrument; CAST: Cognitive assessment screening test; CCCE: Cross-cultural cognitive examination; CCSE: Cognitive capacity screening examination; CERAD: Consortium to Establish a registry for Alzheimer's Disease; CSF: Cerebrospinal fluid; CSI-D: Community screening instrument for dementia; CT: Computed tomography; DaTSCAN: DaT, 123 I-FP-CT; DLB: Dementia with Lewy Bodies; DSM: Diagnostic and statistical manual of mental disorders; DTA: Diagnostic test accuracy; FDG-PET: PET using 2-Fluro-deoxy D-glucose; FTD: Fronto-temporal dementia; ICD-10: International statistical classification of diseases and related health problems, 10th revision; IQCODE: Informant questionnaire on cognitive decline in the elderly; IST: Isaacs set test; M@T: Memory alteration test; MCI: Mild cognitive impairment; MDRS: Mattis dementia rating scale; MFI: Mental function index; MIBG: 123 I-metaiodobenzylguanidine scintigraphy; MIS: Memory impairment screen; MMSE: Mini-mental state examination; MOCA: Montreal cognitive assessment; MRI: Magnetic resonance imaging; NCSE: Neurobehavioral cognitive screening examination; NINCDS-ADRDA: National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer's Disease and Related Disorders Association; NINDS-AIREN: National Institute of Neurological Disorders and Stroke and Association Internationale pour la Recherche et l'Enseignement en Neurosciences; NUCCOG: Neuropsychiatry unit cognitive assessment tool; PIB-PET: 11 C-Pittsburgh Compound B- positron emission tomography; Ptau: Phosphorylated tau; RBANS: Repeatable battery for the assessment of neuropsychological status; RUDAS: Rowland universal dementia assessment scale; SPECT: Single photon emission computed tomography; STMS: Short test of mental status; TMT B: Trail making Test Part B; Ttau: Total tau; VaD: Vascular dementia.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

IAR conceived the review, participated in its design, development and coordination, and helped draft the manuscript. OS participated in the assessment of DTA reviews. IS designed the search strategies and helped draft the manuscript. XB participated in the design of this study and helped draft the manuscript. ES conceived the study, and helped draft the manuscript. PA conceived the review, participated in its design, development and coordination, and helped draft the manuscript. All authors read and approved the final manuscript.

Acknowledgments

The work was supported by the Fundacion Universitaria de Ciencias de la Salud, Bogota- Colombia. Ingrid Arevalo-Rodriguez is a PhD student at the Department of Pediatrics, Obstetrics and Gynecology, and Preventive Medicine of the Universitat Autònoma de Barcelona.

Author details

¹Division of Research, Fundación Universitaria de Ciencias de la Salud, Hospital de San José/Hospital Infantil de San José, Bogotá, DC, Colombia. ²Pediatrics, Obstetrics & Gynecology and Preventive Medicine Department, Universitat Autònoma de Barcelona, Barcelona, Spain. ³Iberoamerican Cochrane Centre, Clinical Epidemiology and Public Health Department, Institute of Biomedical Research (IIB Sant Pau), Spain, CIBER de Epidemiología y Salud Pública (CIBERESP), Barcelona, Spain. ⁴Iberoamerican Cochrane Centre - Institute of Biomedical Research (IIB Sant Pau), CIBER Epidemiología y Salud Pública (CIBERESP), Spain - Universitat Autònoma de Barcelona, Barcelona, Spain. ⁵Department of Neurosciences and Memory Clinic Unit, Fundacion Universitaria de Ciencias de la Salud, Bogotá, Colombia.

Received: 1 June 2014 Accepted: 11 September 2014

Published online: 24 September 2014

References

1. Prince M, Bryce R, Albanese E, Wimo A, Ribeiro W, Ferri CP: **The global prevalence of dementia: a systematic review and metaanalysis.** *Alzheimers Dement* 2013, **9**(1):e62–e75.
2. Wimo A, Winblad B, Jonsson L: **The worldwide societal costs of dementia: Estimates for 2009.** *Alzheimers Dement* 2010, **6**(2):98–103.
3. Davis Daniel HJ, Creavin Sam T, Noel-Storr A, Quinn Terry J, Smailagic N, Hyde C, Brayne C, McShane R, Cullum S: **Neuropsychological tests for the diagnosis of Alzheimer's disease dementia and other dementias: a generic protocol for cross-sectional and delayed-verification studies.** *Cochrane Database Syst Rev* 2013, **3**:CD010460.
4. American Psychiatric Association: *Diagnostic and statistical manual of mental disorders (DSM-IV)*. Washington DC: American Psychiatric Association; 1994.
5. Jack CR Jr, Albert MS, Knopman DS, Mckhann GM, Sperling RA, Carrillo MC, Thies B, Phelps CH: **Introduction to the recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease.** *Alzheimers Dement* 2011, **7**(3):257–262.
6. Mckhann GM, Knopman DS, Chertkow H, Hyman BT, Jack CR Jr, Kawas CH, Klunk WE, Koroshetz WJ, Manly JJ, Mayeux R, Mohs RC, Morris JC, Rossor MN, Scheltens P, Carrillo MC, Thies B, Weintraub S, Phelps CH: **The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease.** *Alzheimers Dement* 2011, **7**(3):263–269.
7. Feldman HH, Jacova C, Robillard A, Garcia A, Chow T, Borrie M, Schipper HM, Blair M, Kertesz A, Chertkow H: **Diagnosis and treatment of dementia: 2. Diagnosis.** *CMAJ* 2008, **178**(7):825–836.
8. Dubois B, Feldman HH, Jacova C, Cummings JL, Dekosky ST, Barberger-Gateau P, Delacourte A, Frisoni G, Fox NC, Galasko D, Gauthier S, Hampel H, Jicha GA, Meguro K, O'Brien J, Pasquier F, Robert P, Rossor M, Salloway S, Sarazin M, de Souza LC, Stern Y, Visser PJ, Scheltens P: **Revising the definition of Alzheimer's disease: a new lexicon.** *Lancet Neurol* 2010, **9**(11):1118–1127.
9. Mak E, Su L, Williams GB, O'Brien JT: **Neuroimaging characteristics of dementia with Lewy bodies.** *Alzheimers Res Ther* 2014, **6**(2):18.
10. Risacher SL, Saykin AJ: **Neuroimaging biomarkers of neurodegenerative diseases and dementia.** *Semin Neurol* 2013, **33**(4):386–416.
11. Reitsma JR AWS, Whiting P, Vlassov W, Leeflang MMG, Deeks JJ: **Chapter 9: Assessing Methodological Quality.** In *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 1.0*. Edited by Deeks JB PM, Gatsonis C: The Cochrane Collaboration; 2009:1–27. Available from: <http://srdta.cochrane.org/>.
12. Arevalo-Rodriguez I, Pedraza OL, Rodriguez A, Sanchez E, Gich I, Sola I, Bonfill X, Alonso-Coello P: **Alzheimer's disease dementia guidelines for diagnostic testing: a systematic review.** *Am J Alzheimers Dis Other Dement* 2013, **28**(2):111–119.
13. Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J: **The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews.** *BMC Med Res Methodol* 2003, **3**:25.
14. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, Leeflang MM, Sterne JA, Bossuyt PM: **QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies.** *Ann Intern Med* 2011, **155**(8):529–536.

15. Shea BJ GJ, Wells GA, Boers M, Andersson N, Hamel C, Porter A, Tugwell P, Moher D, Bouter L: **Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews.** *BMC Med Res Methodol* 2007, **7**:10.
16. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gotzsche PC, Ioannidis JP, Clarke M, Devereaux PJ, Kleijnen J, Moher D: **The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration.** *BMJ (Clinical Research Ed)* 2009, **339**:b2700.
17. Appels BA, Scherder E: **The diagnostic accuracy of dementia-screening instruments with an administration time of 10 to 45 minutes for use in secondary care: a systematic review.** *Am J Alzheimers Dis Other Dement* 2010, **25**(4):301–316.
18. Beynon R, Sterne JA, Wilcock G, Likeman M, Harbord RM, Astin M, Burke M, Bessell A, Ben-Shlomo Y, Hawkins J, Hollingworth W, Whiting P: **Is MRI better than CT for detecting a vascular component to dementia? A systematic review and meta-analysis.** *BMC Neurol* 2012, **12**:33.
19. Bloudek LM, Spackman DE, Blankenburg M, Sullivan SD: **Review and meta-analysis of biomarkers and diagnostic imaging in Alzheimer's disease.** *J Alzheimers Dis* 2011, **26**(4):627–645.
20. Carnero-Pardo C: **Systematic review of the value of positron emission tomography in the diagnosis of Alzheimer's disease.** *Rev Neurol* 2003, **37**(9):860–870.
21. Crawford S, Whitnall L, Robertson J, Evans JJ: **A systematic review of the accuracy and clinical utility of the Addenbrooke's Cognitive Examination and the Addenbrooke's Cognitive Examination-Revised in the diagnosis of dementia.** *Int J Geriatr Psychiatry* 2012, **27**(7):659–669.
22. Dougall NJ, Bruggink S, Ebmeier KP: **Systematic review of the diagnostic accuracy of 99mTc-HMPAO-SPECT in dementia.** *Am J Geriatr Psychiatry* 2004, **12**(6):554–570.
23. Ehreke L, Luppá M, König HH, Riedel-Heller SG: **Is the clock drawing test a screening tool for the diagnosis of mild cognitive impairment? A systematic review.** *Int Psychogeriatr* 2010, **22**(1):56–63.
24. Ferrante D: **SPECT for the Diagnosis and Assessment of Dementia and Alzheimer's Disease.** In *Ciudad de Buenos Aires: Institute for Clinical Effectiveness and Health Policy (IECS)*; 2004.
25. Lischka AR, Mendelsohn M, Overend T, Forbes D: **A systematic review of screening tools for predicting the development of dementia.** *Can J Aging* 2012, **31**(3):295–311.
26. Lonie JA, Tierney KM, Ebmeier KP: **Screening for mild cognitive impairment: a systematic review.** *Int J Geriatr Psychiatry* 2009, **24**(9):902–915.
27. Matchar DB, Kulasingam SL, McCrory DC, Patwardhan MB, Rutschmann OT, Samsa GP, Schmechel DE: **Use of positron emission tomography and other neuroimaging techniques in the diagnosis and management of Alzheimer's disease and dementia.** In Rockville: Agency for Healthcare Research and Quality (AHRQ); 2001.
28. Mitchell AJ: **CSF phosphorylated tau in the diagnosis and prognosis of mild cognitive impairment and Alzheimer's disease: a meta-analysis of 51 studies.** *J Neurol Neurosurg Psychiatry* 2009, **80**(9):966–975.
29. Mitchell AJ: **A meta-analysis of the accuracy of the mini-mental state examination in the detection of dementia and mild cognitive impairment.** *J Psychiatr Res* 2009, **43**(4):411–431.
30. Mitchell AJ, Malladi S: **Screening and case finding tools for the detection of dementia. Part I: evidence-based meta-analysis of multidomain tests.** *Am J Geriatr Psychiatry* 2010, **18**(9):759–782.
31. Mitchell AJ, Malladi S: **Screening and case-finding tools for the detection of dementia. Part II: evidence-based meta-analysis of single-domain tests.** *Am J Geriatr Psychiatry* 2010, **18**(9):783–800.
32. Monge-Argiles JA, Sanchez-Paya J, Munoz-Ruiz C, Pampliega-Perez A, Montoya-Gutierrez J, Leiva-Santana C: **Biomarkers in the cerebrospinal fluid of patients with mild cognitive impairment: a meta-analysis of their predictive capacity for the diagnosis of Alzheimer's disease.** *Rev Neurol* 2010, **50**(4):193–200.
33. Papanthasiou ND, Boutsidiadis A, Dickson J, Bomanji JB: **Diagnostic accuracy of (1)(2)(3)-FP-CIT (DaTSCAN) in dementia with Lewy bodies: a meta-analysis of published studies.** *Parkinsonism Relat Disord* 2012, **18**(3):225–229.
34. Patwardhan MB, McCrory DC, Matchar DB, Samsa GP, Rutschmann OT: **Alzheimer disease: operating characteristics of PET—a meta-analysis.** *Radiology* 2004, **231**(1):73–80.
35. Treglia G, Cason E, Giordano A: **Diagnostic performance of myocardial innervation imaging using MIBG scintigraphy in differential diagnosis between dementia with Lewy bodies and other dementias: a systematic review and a meta-analysis.** *J Neuroimaging* 2012, **22**(2):111–117.

Arevalo-Rodriguez et al. *BMC Neurology* 2014, **14**:183
<http://www.biomedcentral.com/1471-2377/14/183>

Page 8 of 8

36. van Harten AC, Kester MI, Visser PJ, Blankenstein MA, Pijnenburg YA, van der Flier WM, Scheltens P: **Tau and p-tau as CSF biomarkers in dementia: a meta-analysis.** *Clin Chem Lab Med* 2011, **49**(3):353–366.
37. van Rossum IA, Vos S, Handels R, Visser PJ: **Biomarkers as predictors for conversion from mild cognitive impairment to Alzheimer-type dementia: implications for trial design.** *J Alzheimers Dis* 2010, **20**(3):881–891.
38. Yuan Y, Gu ZX, Wei WS: **Fluorodeoxyglucose-positron-emission tomography, single-photon emission tomography, and structural MR imaging for prediction of rapid conversion to Alzheimer disease in patients with mild cognitive impairment: a meta-analysis.** *AJNR Am J Neuroradiol* 2009, **30**(2):404–410.
39. Zhang S, Han D, Tan X, Feng J, Guo Y, Ding Y: **Diagnostic accuracy of 18F-FDG and 11 C-PIB-PET for prediction of short-term conversion to Alzheimer's disease in subjects with mild cognitive impairment.** *Int J Clin Pract* 2012, **66**(2):185–198.
40. Yeo JMLX, Khan Z, Pal S: **Systematic review of the diagnostic utility of SPECT imaging in dementia.** *Eur Arch Psychiatry Clin Neurosci* 2013, **263**(7):539–552.
41. Sackett D, Haynes RB: **The Architecture of Diagnostic Research.** In *The Evidence Base of Clinical Diagnosis*. Edited by Knottnerus J. London: BMJ Publisher; 2002.
42. Birks J, Flicker L: **Donepezil for mild cognitive impairment.** *Cochrane Database Syst Rev* 2006, **3**:CD006104.
43. Martin M, Clare L, Altgassen Anne M, Cameron Michelle H, Zehnder F: **Cognition-based interventions for healthy older people and people with mild cognitive impairment.** *Cochrane Database Syst Rev* 2011, **1**:CD006220.
44. Russ Tom C, Morling Joanne R: **Cholinesterase inhibitors for mild cognitive impairment.** *Cochrane Database Syst Rev* 2012, **9**:CD009132.
45. Yue J, Dong Bi R, Lin X, Yang M, Wu Hong M, Wu T: **Huperzine A for mild cognitive impairment.** *Cochrane Database Syst Rev* 2012, **12**:CD008827.
46. Tatsioni A, Zarin DA, Aronson N, Samson DJ, Flamm CR, Schmid C, Lau J: **Challenges in systematic reviews of diagnostic technologies.** *Ann Intern Med* 2005, **142**(12 Pt 2):1048–1055.
47. Chalmers I, Bracken MB, Djulbegovic B, Garattini S, Grant J, Gülmezoglu AM, Howells DW, Ioannidis JPA, Oliver S: **How to increase value and reduce waste when research priorities are set.** *Lancet* 2014, **383**(9912):156–165.
48. Macaskill PG C, Deeks JJ, Harbord RM, Takwoingi Y: **Chapter 10: Analysing and presenting results.** In *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 1.0*. Edited by Deeks JB PM, Gatsonis C: The Cochrane Collaboration; 2010:1–61. Available from: <http://srdta.cochrane.org/>.
49. Reitsma JB, Glas AS, Rutjes AW, Scholten RJ, Bossuyt PM, Zwinderman AH: **Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews.** *J Clin Epidemiol* 2005, **58**(10):982–990.
50. Harbord RM, Whiting P, Sterne JA, Egger M, Deeks JJ, Shang A, Bachmann LM: **An empirical comparison of methods for meta-analysis of diagnostic accuracy showed hierarchical models are necessary.** *J Clin Epidemiol* 2008, **61**(11):1095–1103.
51. Hayden AMP, Irwig L, Bossuyt P: **Appropriate statistical methods are required to assess diagnostic tests for replacement, add-on, and triage.** *J Clin Epidemiol* 2010, **63**(8):883–891.
52. Whiting P, Rutjes AWS, Reitsma JB, Glas AS, Bossuyt PM, Kleijnen J: **Sources of variation and bias in studies of diagnostic Accuracy: a systematic review.** *Ann Intern Med* 2004, **140**(3):189–202.
53. Leeflang MM, Deeks JJ, Rutjes AW, Reitsma JB, Bossuyt PM: **Bivariate meta-analysis of predictive values of diagnostic tests can be an alternative to bivariate meta-analysis of sensitivity and specificity.** *J Clin Epidemiol* 2012, **65**(10):1088–1097.
54. Deeks JJ, Macaskill P, Irwig L: **The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed.** *J Clin Epidemiol* 2005, **58**(9):882–893.
55. Noel-Storr AH, McCleery JM, Richard E, Ritchie CW, Flicker L, Cullum SJ, Davis D, Quinn TJ, Hyde C, Rutjes AW, Smailagic N, Marcus S, Black S, Blennow K, Brayne C, Fiorivanti M, Johnson JK, Kopke S, Schneider LS, Simmons A, Mattsson N, Zetterberg H, Bossuyt PM, Wilcock G, McShane R: **Reporting standards for studies of diagnostic test accuracy in dementia: The STARDDem Initiative.** *Neurology* 2014, **83**(4):364–373.

doi:10.1186/s12883-014-0183-2

Cite this article as: Arevalo-Rodriguez et al.: Diagnostic tools for alzheimer's disease dementia and other dementias: an overview of diagnostic test accuracy (DTA) systematic reviews. *BMC Neurology* 2014 **14**:183.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit

