



EFECTIVIDAD DE LA RESONANCIA MAGNÉTICA CEREBRAL EN EL DIAGNÓSTICO, PRONÓSTICO Y CARACTERIZACIÓN DE LAS DEMENCIAS: UNA REVISIÓN SISTEMÁTICA.

Tesina Fin de Máster

Alumna: Ana M^a Ruiz Tornero

Tutoras: Esther E. García Carpintero. Begoña Rodríguez Ortiz de Salazar.

Centro de realización: Agencia de Evaluación de Tecnologías Sanitarias (Instituto de Salud Carlos III). Consejo Superior de Investigaciones Científicas.

Dirección: C/ de Monforte de Lemos 5, 28029. Madrid.

ÍNDICE

PREFACIO.....	2
RESUMEN Y PALABRAS CLAVE	3
LISTA DE TABLAS Y FIGURAS.....	4
LISTA DE ABREVIATURAS.....	5
INTRODUCCIÓN	6
MATERIAL Y MÉTODOS	8
RESULTADOS	10
1. Conversión de deterioro cognitivo leve a enfermedad de Alzheimer	13
1.1. Hipocampo	13
1.2. Lóbulo temporal.....	14
1.3. Sustancia gris.....	15
1.4. Combinación de varias estructuras cerebrales	15
2. Conversión de deterioro cognitivo leve a otros tipos de demencia	15
DISCUSIÓN.....	18
CONCLUSIONES	22
BIBLIOGRAFÍA.....	23
TABLAS Y FIGURAS.....	28
ACTIVIDAD DESARROLLADA E IMPLICACIÓN DE LA ESTUDIANTE.....	46
ANEXOS	47

PREFACIO

Este trabajo ha sido realizado entre diciembre de 2021 y julio de 2022, y se presenta como Tesina Fin de Máster del Máster de Salud Pública de la Escuela Nacional de Sanidad.

Quiero dar las gracias al Instituto de Salud Carlos III por transmitirme la formación necesaria para elaborar el presente trabajo, y en especial a mis tutoras Esther y Begoña, y a mi mentora M^º Isabel Portillo por compartir su experiencia y conocimientos conmigo. También a mi tutora Alicia Díaz Redondo del hospital Gregorio Marañón por su disponibilidad y cercanía. Finalmente, a mi familia y amigos por su paciencia y escucha en los malos momentos y por su compañía en los buenos, sin ellos nada de esto sería posible.

RESUMEN Y PALABRAS CLAVE

Introducción: La demencia es una enfermedad neurodegenerativa que cursa con el deterioro progresivo de una o varias áreas cognitivas, siendo la enfermedad de Alzheimer (EA) la causa más frecuente. La resonancia magnética (RM) cerebral se utiliza para descartar lesiones subyacentes, si bien también podría servir para identificar patrones de atrofia característicos de determinados tipos de demencia. El objetivo de este trabajo es evaluar la evidencia científica sobre la efectividad de la RM cerebral en el diagnóstico de las demencias, así como su capacidad para el diagnóstico precoz en pacientes con deterioro cognitivo leve (DCL).

Material y métodos: Se seleccionaron estudios primarios a partir de revisiones sistemáticas entre los años 2011 y 2021. Se siguieron los criterios PRISMA del año 2020. Se utilizó la herramienta QUADAS-2 para analizar la calidad de los estudios, y la certeza de la evidencia se evaluó con la herramienta GRADE. Se extrajo la información más relevante de los artículos, incluyendo los valores de sensibilidad y especificidad. Además, se efectuó un metanálisis con los estudios que aportaban datos suficientes (n=11).

Resultados: Se identificó una revisión sistemática de alta calidad de la que se extrajeron 15 estudios, seleccionando otros 8 estudios a partir de una actualización de la búsqueda, con una selección final de 23 artículos para el análisis. Al analizar la calidad de los estudios, el apartado de selección de pacientes y la prueba índice presentaron en general alta probabilidad de sesgo, y las categorías de prueba de referencia y flujo y tiempos baja probabilidad de sesgo. La certeza de la evidencia con la herramienta GRADE fue muy baja. Para el hipocampo, la sensibilidad media hallada fue de 0,62 (IC95% 0,48 – 0,79) y la especificidad media de 0,70 (IC95% 0,55 – 0,80). Para el lóbulo temporal, la media de sensibilidad fue de 0,65 (rango 0,45) y la especificidad de 0,69 (rango 0,32). Para el resto de áreas cerebrales, existían escasos estudios disponibles y homogéneos para establecer valores de sensibilidad y especificidad globales.

Conclusiones: No existe suficiente evidencia para recomendar la RM cerebral de forma rutinaria para el diagnóstico precoz de demencia en pacientes con DCL, con unos valores de sensibilidad y especificidad obtenidos en el metanálisis de 0,67 y 0,68 respectivamente, siendo la certeza de la evidencia muy baja. En un futuro, sería de utilidad la realización de estudios adicionales con mayor homogeneidad en su metodología para aportar mayor calidad en el momento de comparar sus resultados.

PALABRAS CLAVE: Demencia; Enfermedad de Alzheimer; Deterioro cognitivo leve; Resonancia magnética; Diagnóstico precoz; Precisión diagnóstica.

LISTA DE TABLAS Y FIGURAS

Tabla 1: Estrategia de búsqueda.....	28
Tabla 2: Características principales de los estudios incluidos.....	29
Tabla 3: Valores de sensibilidad y especificidad de los estudios incluidos.....	30
Tabla 4: Tabla de resultados del QUADAS-2.....	38
Tabla 5: Tabla de evidencia de GRADE.....	39
Figura 1: Diagrama de flujo PRISMA sobre la selección de estudios.....	40
Figura 2: Gráfica QUADAS-2 de los estudios incluidos.....	41
Figura 3: Metanálisis de los estudios incluidos con datos disponibles.....	42
Figura 4: Curva ROC	43
Figura 5: Metanálisis por subgrupo.....	44
Figura 6: Curva ROC del subgrupo	45

LISTA DE ABREVIATURAS

Siglas	Aclaración terminológica
APP	Afasia primaria progresiva
APPnf	Afasia primaria progresiva no fluente
APPse	Afasia primaria progresiva semántica
DCL	Deterioro cognitivo leve
DCLewy	Demencia por cuerpos de Lewy
DCS	Deterioro cognitivo subjetivo
DFT	Demencia frontotemporal
DFTco	Demencia frontotemporal variante conductual
DSM-IV	Diagnostic and Statistical Manual of Mental Disorders-IV
DVa	Demencia vascular
EA	Enfermedad de Alzheimer
MMSE	Mini-Mental State Examination
NIA-AA	National Institute on Aging-Alzheimer's Association
NINCDS/ADRDA	National Institute of Neurological and Communicative Disorders and Stroke/Alzheimer's Disease and Related Disorders Association
RM	Resonancia magnética
TC	Tomografía computarizada

INTRODUCCIÓN

La demencia es una enfermedad neurodegenerativa que cursa con el deterioro progresivo de una o varias áreas cognitivas ⁽¹⁾, con el consecuente desarrollo de discapacidad y dependencia para la realización de las actividades de la vida diaria ⁽²⁾. En concreto, la enfermedad de Alzheimer (EA) es la primera causa de demencia neurodegenerativa a nivel mundial, cursando con dos posibles formas de presentación: una típica con deterioro amnésico y disfunción ejecutiva iniciales; y otra atípica con deterioro de otras funciones cognitivas (lenguaje, visual, capacidad ejecutiva) antes de la presentación del deterioro amnésico ⁽²⁾. Tras la EA, la demencia vascular (DVA) es la más frecuente, seguida de la demencia por cuerpos de Lewy (DCLewy) y, con menor frecuencia, la demencia frontotemporal (DFT) ⁽³⁾.

La EA y otras demencias son la séptima causa de muerte, existiendo más de 55 millones de personas diagnosticadas en todo el mundo, sin contar con los casos no diagnosticados por la falta de conocimiento sobre los signos y síntomas de la enfermedad, los sesgos culturales o geográficos o la falta de accesibilidad a recursos y profesionales capacitados ⁽⁴⁾. La prevalencia a nivel mundial en el año 2019 se estimaba en 7,7% de la población entre 75 y 89 años de edad, según datos de Global Burden of Disease del *Institute for Health Metrics and Evaluation* ⁽⁵⁾. El diagnóstico precoz es extremadamente importante en cuanto al pronóstico, ya que puede contribuir a iniciar un tratamiento más temprano, con el objetivo de ralentizar la progresión de la enfermedad, obteniendo mejores resultados terapéuticos y una mejor calidad de vida para el paciente y su familia ⁽⁶⁾.

El concepto de EA preclínica surgió a finales del siglo XX como referencia a la aparición de lesiones cerebrales sugerentes de EA en pacientes sin deterioro cognitivo demostrable ⁽⁷⁾. Aun así, existen a día de hoy evidentes dificultades para detectar este tipo de pacientes, dada la ausencia de síntomas de enfermedad y por tanto la no asistencia a una consulta médica por dicho motivo. Por otro lado, la entidad conocida como deterioro cognitivo leve (DCL) se caracteriza por la afectación progresiva de uno o más dominios cognitivos, encontrándose un empeoramiento mayor de lo que se esperaría para la edad y el nivel educativo del paciente, pero sin existir en ningún momento dependencia para actividades funcionales de la vida diaria ⁽⁸⁾, a diferencia de lo que ocurriría en un paciente con demencia.

La resonancia magnética (RM) cerebral es una técnica de neuroimagen utilizada con frecuencia en la práctica clínica. La potencia del campo magnético que se utiliza con mayor frecuencia es de 1,5 teslas, si bien la potencia de 3 teslas se encuentra disponible en ocasiones en clínicas con

mayor especialización y centros de investigación ⁽⁹⁾. Frente a otras técnicas de imagen presenta varias ventajas, es una técnica no invasiva, que presenta mayor resolución de la imagen, en comparación con la Tomografía Computarizada (TC) cerebral y ausencia de radiación ionizante. La RM cerebral puede ofrecer diferentes técnicas frente a la TC que contribuyen a un mejor diagnóstico diferencial en las demencias, por lo que su utilidad no sería únicamente descartar lesiones subyacentes ⁽¹⁰⁾. Entre sus ventajas actuales, se encuentran el hallazgo de lesiones cerebrales, la posibilidad de determinar patrones de atrofia y la identificación de la carga vascular ⁽¹¹⁾. Además, existen signos demostrables en la RM que pueden detectar de forma precoz un caso de EA, como son el menor volumen del hipocampo, de regiones del lóbulo temporal medial, de la corteza entorrinal o del volumen de sustancia gris en el lóbulo temporal, así como atrofia del córtex prefrontal ^(7,11).

Actualmente, la RM cerebral no posee un papel bien definido en la práctica clínica, considerándose una prueba que en muchos centros es utilizada únicamente para descartar lesiones subyacentes. Sin embargo, dados los avances acontecidos las últimas décadas en técnicas de imagen (y en concreto en la RM) ⁽¹²⁾, creemos necesaria una revisión sistemática de la literatura para sintetizar la evidencia más reciente sobre otras potenciales utilidades en la práctica clínica habitual hoy en día. Por ello, el objetivo principal de esta revisión es:

- Evaluar la evidencia científica sobre la efectividad de la RM en el diagnóstico de las demencias.

Y los objetivos secundarios:

- Evaluar la efectividad de la RM en cada tipo de demencia: demencia frontotemporal, demencia por cuerpos de Lewy o demencia vascular.
- Identificar patrones de imagen de RM asociados al diagnóstico de demencia.
- Cuantificar la sensibilidad y especificidad de la RM cerebral para el diagnóstico precoz de demencia en pacientes con deterioro cognitivo leve.

MATERIAL Y MÉTODOS

El diseño del estudio realizado consiste en una revisión sistemática de la literatura. La Tabla 1 muestra la estrategia de búsqueda. Se realizó una búsqueda en las bases de datos Medline y EMBASE, escogiendo revisiones sistemáticas que trataran sobre el tema de RM cerebral en el diagnóstico de demencia. Entre éstas, se identificó una revisión sistemática reciente de alta calidad de Cochrane, y se incluyeron los estudios primarios que cumplían los criterios de inclusión, identificando además estudios primarios posteriores a la fecha de dicha revisión sistemática con una actualización de la búsqueda bibliográfica. Como criterios de inclusión, se establecieron los estudios realizados a partir del año 2011 y hasta diciembre de 2021, en idioma inglés o castellano. El motivo de escoger estudios a partir de dicho año se fundó en la prioridad de actualizar la última guía de práctica clínica sobre demencias existente en España realizada por la Agencia de Evaluación de Tecnologías Sanitarias, que data del año 2011. Respecto al diseño, se incluyeron estudios transversales, de cohortes o ensayos clínicos. La muestra estaba compuesta de pacientes con sospecha de demencia o deterioro cognitivo leve de cualquier edad y sexo o bien pacientes con diagnóstico clínico ya establecido de cualquier tipo de demencia. Además, el objetivo principal de los estudios estaba definido por la valoración de la precisión diagnóstica de la RM cerebral para el diagnóstico de demencia, aportando datos concretos sobre sensibilidad y especificidad. Se excluyeron artículos que tuvieran un diseño de casos y controles, así como estudios con metodología cualitativa. También se excluyeron artículos cuyo objetivo principal no fuera la valoración de la RM cerebral en sí, así como estudios cuyo objetivo principal fuera la investigación de la metodología de inteligencia artificial y aprendizaje automático. Inicialmente, se realizó una primera revisión por título y resumen, seleccionando los estudios a incluir en la revisión. En un segundo tiempo, se realizó un análisis de los estudios por texto completo seleccionando los que definitivamente cumplían los criterios de inclusión para la revisión sistemática. Este proceso fue realizado por tres investigadoras de forma independiente. Para realizar una colaboración de manera adecuada y sencilla entre las diferentes investigadoras, se utilizó la plataforma Rayyan, donde se compartieron los artículos seleccionados, se etiquetaron como incluidos o excluidos y se realizaron anotaciones en los que se encontraban dudosos. Cuando existían conflictos sobre la inclusión o exclusión de algún artículo, se acordaba una reunión deliberativa entre todo el equipo investigador para consensuar la decisión final.

La información más relevante se recogió en tablas realizadas en Microsoft Excel a partir de los estudios primarios seleccionados. Los valores medios de sensibilidad y especificidad y sus intervalos de confianza al 95% se calcularon con este mismo programa. Para cada estudio analizado, se recogieron datos sobre: autor y título; lugar de realización; muestra; edad media, porcentaje de mujeres y puntuación media del test minimal (MMSE) de los participantes; diseño del estudio; tiempo de seguimiento; número de pacientes con evolución a demencia; prueba de referencia; sensibilidad; especificidad; resumen de principales resultados.

La evaluación del riesgo de sesgo se realizó con la herramienta QUADAS-2 ⁽¹³⁾. Para facilitar la síntesis de estos resultados, se utilizó el programa informático Review Manager 5.4. La certeza de la evidencia se evaluó con la herramienta *Grading of Recommendations Assessment Development and Evaluation* (GRADE), que aporta información sobre la calidad de la evidencia y permite establecer una graduación en la fuerza de las recomendaciones durante la elaboración de guías de práctica clínica, revisiones sistemáticas o evaluación de tecnologías sanitarias ^(14,15).

Para efectuar el metanálisis se usó el programa informático EasyMeta con la importación de los datos de las tablas 2x2 que se encontraban disponibles en los artículos, desde una hoja de cálculo de Microsoft Excel. Respecto a la metodología del metanálisis, se calcularon los indicadores agregados según cada variable resultado considerada, con sus intervalos de confianza al 95%. El nivel de heterogeneidad se evaluó mediante el estadístico I^2 . Un $I^2 < 30\%$ indica una heterogeneidad baja; entre $30\% \leq I^2 \leq 50\%$ indica una heterogeneidad moderada; y un valor de $I^2 > 50\%$ una heterogeneidad alta. Si la heterogeneidad tanto estadística (calculada mediante un test de χ^2) como clínica, resultasen estadísticamente significativa o sustancial, se empleó un modelo de efectos aleatorios. También se realizó un análisis de subgrupo en aquellos estudios que habían realizado medición del volumen completo de hipocampo con medidas automatizadas, por ser considerado éste el grupo más homogéneo. Se calcularon además las curvas ROC con los datos obtenidos.

La presente revisión siguió los criterios *Preferred Reporting Items for Systematic Reviews and Meta-Analyses* (PRISMA) del año 2020 ⁽¹⁶⁾.

RESULTADOS

Inicialmente se encontraron 414 artículos que cumplían los criterios de inclusión. Tras leer el título y el resumen, y hacer una revisión cualitativa de las bibliografías de los artículos, 81 artículos cumplían los criterios de inclusión y el resto fueron excluidos. De ellos, se desestimaron 15 por no aportar datos de sensibilidad ni especificidad; 10 por no presentar un diseño acorde con los criterios de inclusión (siendo 7 estudios de casos y controles y 3 revisiones); 20 por ser anteriores al año 2011; 1 por encontrarse duplicado; y 12 por no cumplir el objetivo de nuestro trabajo. Finalmente, se obtuvieron 23 artículos originales que se incluyeron para la revisión sistemática (referencias 18-28 y 30-41). De ellos, 15 artículos que cumplían los criterios de inclusión se han extraído de una revisión sistemática previa de Cochrane⁽¹⁷⁾, y 8 fueron añadidos de forma adicional como estudios primarios al actualizar la búsqueda. Entre los artículos seleccionados, 20 estudiaban la efectividad de la RM cerebral para el diagnóstico precoz de EA en pacientes con DCL, 1 investigaba la efectividad de la RM en EA y otro tipo de demencias, y 2 solamente en otro tipo de demencias (uno de ellos en DCLewy y otro en DFT). La lista de verificación PRISMA se encuentra disponible en el anexo I, y el diagrama PRISMA sobre la selección de estudios se muestra en la figura 1. En la Tabla 2 se muestran las principales características de los artículos analizados, y en la Tabla 3 los valores de sensibilidad y especificidad aportados en cada uno de los estudios.

Se analizó la calidad de los estudios incluidos con la herramienta QUADAS-2⁽¹³⁾. La Tabla 4 muestra los resultados del riesgo de sesgo para cada estudio, y la figura 2 los resume de forma gráfica. En el ítem de selección de pacientes, solamente un estudio⁽¹⁸⁾ presentó baja probabilidad de sesgo, 3 estudios⁽¹⁹⁻²¹⁾ mostraron probabilidad incierta, y el resto de estudios mostraron alta probabilidad de sesgo. La elevada probabilidad de sesgo estuvo en relación con la falta de aleatorización de los pacientes incluidos en los estudios. En cuanto a la prueba índice, 3 estudios mostraron probabilidad baja de sesgo⁽²²⁻²⁴⁾, 6 estudios probabilidad incierta, y el resto de estudios (n=14) mostraron probabilidad alta. En este caso, la alta probabilidad estuvo motivada por el conocimiento del diagnóstico de la enfermedad con la prueba de referencia a la hora de interpretar la prueba índice. En cuanto a la prueba de referencia, se consideró como tal el diagnóstico con los criterios clínicos de demencia, utilizando la mayoría de los estudios los criterios del *National Institute on Aging-Alzheimer's Association (NIA-AA)*, del *Diagnostic and Statistical Manual of Mental Disorders-IV (DSM-IV)* o los del *National Institute of Neurological and Communicative Disorders and Stroke/Alzheimer's Disease and Related Disorders Association*

(NINCDS/ADRDA), considerándose todos ellos válidos. La mayoría de estudios en este apartado presentaron probabilidad baja de sesgo, mientras que un estudio presentó probabilidad alta ⁽²¹⁾ (por no utilizar criterios claramente establecidos para el diagnóstico de demencia, sino un empeoramiento en actividades básicas o instrumentales según una escala), y nueve estudios mostraron probabilidad incierta. En cuanto al flujo y tiempos, la mayoría de estudios también mostraron probabilidad baja de sesgo, 4 estudios presentaron probabilidad incierta, y ningún estudio mostró probabilidad elevada de sesgo. Respecto a la preocupación por la aplicabilidad de los resultados, todos los estudios mostraron baja preocupación tanto en la selección de pacientes, como en la prueba índice y en la prueba de referencia, con la excepción de un estudio que mostró probabilidad incierta en cuanto a la prueba de referencia ⁽²¹⁾ por la falta de utilización de una prueba de referencia validada. Como dato a añadir respecto a la calidad de los estudios, es preciso constatar que no todos aportaban los intervalos de confianza de los valores de sensibilidad y especificidad.

La certeza de la evidencia se evaluó con la herramienta GRADE ⁽¹⁴⁾. Se obtuvo una certeza muy baja, según se muestra en la Tabla 5. Este grado de certeza fue debido principalmente a sesgos en la selección de pacientes (falta de aleatorización) y en la prueba índice (conocimiento del diagnóstico con la prueba de referencia en el momento de realizar la prueba índice), así como por inconsistencia entre algunos estudios e imprecisión, mostrando intervalos de confianza muy amplios en algunos casos.

Respecto al área cerebral concreta investigada, la mayoría de los estudios (n=13) se enfocaban en el volumen del hipocampo, valorando en conjunto uno de ellos el córtex entorrinal ⁽²⁵⁾ y otro el lóbulo temporal medial y el ventrículo lateral ⁽²⁶⁾; 3 estudios medían exclusivamente la atrofia del lóbulo temporal; un estudio medía el volumen de sustancia gris; 1 estudio valoraba la pérdida del signo de la cola de la golondrina en la demencia por cuerpos de Lewy; 1 estudio evaluaba el volumen de la amígdala; 1 estudio valoraba la atrofia frontal aislada en la demencia frontotemporal; y el resto de estudios (n=3) evaluaban varias estructuras cerebrales y la combinación entre ellas. La muestra total analizada en los 23 estudios fue de 5761 sujetos. La progresión media a demencia en los pacientes con DCL analizados fue del 42%. En los estudios que detallaban de forma separada la edad de los participantes con DCL estable y progresivo, la edad media de los pacientes con DCL estable fue ligeramente menor respecto al DCL progresivo: 71,6 años en DCL estable (rango 11,2 años), frente a 73,5 años en DCL progresivo (rango 6,3 años). La sensibilidad mínima reportada fue de 0,28 (en una de las cohortes del estudio de Prestia et al. ⁽²⁴⁾ con la medición semi-automatizada del volumen total del hipocampo) y la

máxima de 0,89 (en el estudio de Nesteruk et al., utilizando la combinación de varias estructuras cerebrales⁽²⁷⁾). La especificidad mínima reportada fue de 0,42 (en el estudio de Gaser et al, con la medición automatizada del volumen de hipocampo derecho⁽²⁸⁾), y la máxima de 0,94 (en una de las cohortes del estudio de Prestia et al.,⁽²⁴⁾ con la medición semi-automatizada del volumen total del hipocampo). La sensibilidad y especificidad medias de forma global, independientemente de la técnica de medida o la región cerebral analizada, fueron de 0,67 (IC95% 0.57 - 0.80) y 0,72 (IC95% 0.56 - 0.78) respectivamente. La figura 3 muestra el diagrama de bosque o *forest plot* con los estudios de los que se disponía de datos suficientes para hacer un análisis cuantitativo, utilizando un modelo de efectos aleatorios. Se obtuvo una sensibilidad de 0,67 (IC95% 0,58 – 0,74) y una especificidad de 0,68 (IC95% 0,60 – 0,75). La heterogeneidad medida con I^2 fue de 60,73% en la sensibilidad y 78,62% en la especificidad, con una $P=0,005$ para la sensibilidad y $P<0,001$ para la especificidad. Es destacable que en algunos estudios los intervalos de confianza al 95% mostrados fueron muy anchos (por ejemplo, en los estudios de Ferrari et al., Ong et al. o Wood et al.), lo que generó mayor incertidumbre sobre la veracidad de los resultados finales obtenidos. Además, existían algunos estudios que se alejaron más que otros de los valores globales de sensibilidad y especificidad. En cuanto a la primera, el estudio de Prestia et al. mostró una sensibilidad de 0,28 (respecto al 0,67 global), pudiendo ser una de las razones de esta diferencia con el valor global el hecho de que solo se aportaron datos para el metanálisis de una de las dos cohortes analizadas en el estudio, conteniendo una muestra muy pequeña (36 sujetos). Respecto a la especificidad, no se halló tanta diversidad entre los valores individuales de cada estudio, siendo el de Khan et al. el que más difería del valor global, con una especificidad de 0,50 respecto al 0,68 global. Se obtuvo una razón de verosimilitud negativa de 0,45 (IC95% 0,38 – 0,54), y positiva de 2,0 (IC95% 1,69 – 2,42), considerándose ambos valores de utilidad regular (siendo en el caso de la razón de verosimilitud positiva <2 de utilidad mala, 2-5 regular, 5-10 buena y >10 altamente relevante, y en el caso de la negativa >0,5 mala, 0,5-0,2 regular, 0,1-0,2 buena y <0,1 altamente relevante)⁽²⁹⁾. La figura 4 ilustra la curva ROC, mostrando una acumulación de los valores de sensibilidad en torno a 0,7 y del inverso de la especificidad en torno a 0,4.

A continuación, se procede a presentar los resultados encontrados para las diferentes regiones cerebrales expuestas.

1. Conversión de DCL a EA

1.1. Hipocampo

Se analizaron 13 estudios que valoraban el volumen del hipocampo como marcador del riesgo de EA ^(20,21,24–26,28,30–36). La muestra total analizada estaba constituida por 2335 sujetos. La edad media de los participantes fue de 72,2 años (rango 11,2 años) y el porcentaje de mujeres medio fue de 47,6%. La puntuación basal media de deterioro cognitivo según el MMSE fue de 26,8 puntos (rango 7 puntos). Se utilizaron diferentes técnicas de medición del volumen de hipocampo: únicamente manual (en un estudio ⁽³³⁾) o únicamente automatizada (en 10 estudios). Además, en un estudio se reportaron tanto mediciones manuales como automatizadas ⁽²⁶⁾, y en otro estudio se reportaron tanto mediciones automatizadas como semi-automatizadas ⁽²⁴⁾. Para calcular la precisión diagnóstica, en 11 estudios se usó el volumen total del hipocampo (de forma bilateral o bien con la media de ambos lados), y en los otros dos se midió el lado derecho y el izquierdo de forma separada. Independientemente de la medición del volumen de hipocampo de forma uni o bilateral y del método de medida, la sensibilidad media hallada fue de 0,62 (IC95% 0,48 – 0,79) y la especificidad media de 0,70 (IC95% 0,55 – 0,80).

En los estudios que reportaban el volumen total del hipocampo, la sensibilidad mínima reportada fue de 0,28 (en una de las dos cohortes analizadas en el estudio de Prestia et al., utilizando un tipo de medición semi-automatizada) y la máxima de 0,86 (en el estudio de Wood et al., con el uso de medición automatizada). En cuanto a la especificidad, la mínima reportada fue de 0,50 (en el estudio de Khan et al., con medición automatizada) y la máxima de 0,94 (en una de las dos cohortes analizadas en el estudio de Prestia et al., con el uso de un tipo de medición semi-automatizada). La sensibilidad y especificidad medias reportadas en los estudios que utilizaban la medición del volumen de hipocampo de forma bilateral, independientemente del método de medición utilizado, fueron de 0,59 (IC95% 0,48 – 0,79) y 0,71 (IC95% 0,55 – 0,80) respectivamente. En el estudio que realizó la medición del volumen de hipocampo total con métodos manuales ⁽²⁶⁾, la sensibilidad reportada fue de 0,78 y la especificidad de 0,65. En el estudio que reportó datos con medición semi-automatizada ⁽²⁴⁾, la sensibilidad media de las dos cohortes analizadas fue de 0,37 y la especificidad media de 0,85. En los estudios que utilizaron técnicas automatizadas de medida, la sensibilidad media fue de 0,61 (IC95% 0,46 – 0,78) y la especificidad media fue de 0,69 (IC95% 0,55 – 0,81). La figura 5 muestra los resultados del metanálisis realizado en este subgrupo, con los estudios que otorgaban datos para completar la tabla 2x2, obteniendo una sensibilidad de 0,63 (IC95% 0,49 – 0,75) con una heterogeneidad de

I² 67,05% y P=0,006, y una especificidad de 0,70 (IC95% 0,56 – 0,81) con una heterogeneidad de I² 83,82% y P<0,001. De nuevo, en los valores extremos de sensibilidad de los estudios individuales destaca el artículo de Prestia et al. por las razones comentadas previamente, encontrándose valores de especificidad más homogéneos entre los estudios individuales. La figura 6 muestra la curva ROC, donde se observa cómo el valor de sensibilidad se encuentra en torno a 0,7 y el inverso de la especificidad alrededor de 0,35. En el único estudio que evaluó tanto la medición manual como la automatizada ⁽²⁶⁾, la sensibilidad fue mayor con la medición de tipo manual (0,78 vs 0,66), y la especificidad fue mayor con la medición automatizada (0,77 vs 0,65).

En cuanto a la lateralidad, en uno de los dos estudios analizados (de Gaser et al.) el hipocampo derecho mostró mayor sensibilidad que el izquierdo ⁽²⁸⁾, y en el otro estudio (de Eckerström et al.) se observó lo contrario ⁽³³⁾. Respecto a la especificidad, el hipocampo izquierdo mostró mayor precisión en el estudio de Gaser et al. (0,81 vs 0,42), lo que también ocurrió en el estudio de Eckerström et al. respecto a la conversión a cualquier tipo de demencia (0,66 vs 0,57), pero no respecto a la conversión a EA, donde el lado derecho se mostró discretamente más específico que el izquierdo (0,81 vs 0,75 respectivamente).

1.2. Lóbulo temporal

Dentro de los 3 estudios que utilizaron exclusivamente la atrofia del lóbulo temporal como marcador del riesgo de desarrollar EA en pacientes con DCL ^(22,23,37), la muestra total estaba conformada por 1388 sujetos. La edad media fue de 72,9 años, el porcentaje de mujeres medio fue del 48,7% y la puntuación de MMSE basal media fue de 25 puntos. La medición de la atrofia se realizó mayoritariamente de forma manual, aunque en un estudio se comparó con la herramienta prediseñada PredictAD. En otro estudio además se evaluó la atrofia del lóbulo temporal medial junto con el volumen del hipocampo y del ventrículo lateral ⁽²⁶⁾. En este caso se reportaron los datos con medición manual, siendo la sensibilidad de 0,66 y la especificidad de 0,64. La sensibilidad mínima reportada en estos 4 estudios fue de 0,40 y la máxima de 0,85 (con métodos manuales). La especificidad mínima reportada fue de 0,53 y la máxima de 0,85 (también con métodos manuales). La media de sensibilidad hallada entre los 4 estudios fue de 0,65 (rango 0,45) y la especificidad de 0,69 (rango 0,32). En estos artículos los autores no aportaban intervalos de confianza para los valores de sensibilidad y especificidad.

1.3. Sustancia gris

El estudio de Douaud et al. valoró el volumen de sustancia gris como marcador del desarrollo de EA en pacientes con DCL⁽¹⁹⁾. La sensibilidad reportada fue de 0,39 y la especificidad de 0,82 (sin aportar intervalos de confianza), encontrándose que el volumen de sustancia gris era ligeramente más pequeño en pacientes con DCL progresivo frente a estable, pero sin presentar una sensibilidad aceptable para predicción del desarrollo de demencia en pacientes con DCL. La muestra analizada fue de 56 pacientes con DCL. La edad media de los pacientes con DCL estable fue de 69 ± 9 años, mientras que con DCL progresivo fue de 76 ± 6 años. Es destacable que la mayoría de los pacientes con DCL que progresaron a demencia eran varones (77%), encontrándose esta proporción entre sexos equilibrada en los pacientes que se mantuvieron con DCL estable (50%).

1.4. Combinación de varias estructuras cerebrales

El estudio de Nesteruk et al. seleccionó varias estructuras cerebrales: el hipocampo, el córtex entorrinal, el giro cingulado posterior, el giro parahipocampal, el giro temporal superior, medio e inferior y el volumen intracraneal total⁽²⁷⁾. La muestra estaba compuesta por 40 pacientes con DCL, de los cuales 9 (22,5%) progresaron a EA. La sensibilidad y especificidad calculadas para dichos volúmenes fueron de 0,89 y 0,90 respectivamente, siendo el hipocampo derecho e izquierdo, el córtex entorrinal izquierdo y el giro temporal medio derecho las regiones cerebrales que mostraron diferencias estadísticamente significativas para diferenciar los pacientes que progresaron a EA frente a los que no lo hicieron.

Finalmente, el estudio de Ledig et al.⁽³⁸⁾ utilizó diferentes estructuras cerebrales para valorar su precisión en el diagnóstico precoz de EA en pacientes con DCL: la amígdala, el ventrículo lateral, el hipocampo, el área entorrinal, la sustancia gris o el volumen cerebral total. Sobre una muestra de 343 pacientes con DCL, encontró que la región cerebral más sensible para diferenciar el DCL que progresa a EA del estable fue la amígdala (con una sensibilidad de 0,63 y una especificidad de 0,68), y la más específica fue el ventrículo lateral (con una sensibilidad de 0,54 y una especificidad de 0,74).

2. Conversión de DCL a otros tipos de demencia

En el estudio retrospectivo de Orso et al. se analizó la atrofia aislada del lóbulo frontal como marcador de conversión a demencia en pacientes con deterioro conductual leve⁽¹⁸⁾. Se halló en

un subgrupo de 57 pacientes una sensibilidad de 0,77 y una especificidad de 0,93, entre los cuales el 22,8% progresaron a DFT variante conductual. La edad media de la muestra total de pacientes (113 sujetos) fue de $66 \pm 7,2$ años, y también es destacable la menor proporción de mujeres en la muestra total de sujetos (28,3%).

En otro estudio prospectivo con 2 años de seguimiento se valoró la atrofia de la amígdala como marcador del riesgo de desarrollar EA ⁽³⁹⁾. Se analizaron 60 sujetos, padeciendo 34 de ellos DCL de tipo amnésico. La progresión a demencia tipo Alzheimer en este subgrupo de pacientes con DCL fue del 47%. La edad media era similar en los pacientes con DCL estable respecto a progresivo ($68,4 \pm 7,1$ años frente a $69,7 \pm 6,5$ años respectivamente). El porcentaje de mujeres fue mayor en el grupo de pacientes con DCL estable (72,2%) respecto a progresivo (56,2%). La medición de la atrofia de amígdala derecha mostró una sensibilidad de 0,74 y una especificidad de 0,79 para la predicción del desarrollo de EA en pacientes con DCL.

Un estudio valoró la precisión diagnóstica de la pérdida del signo de la cola de la golondrina como marcador para DCLewy ⁽⁴⁰⁾. Este signo consiste en la pérdida del brillo habitual en el tercio posterior de la sustancia negra del mesencéfalo. Se realizó un seguimiento de un año de duración en un grupo de 44 pacientes con diferentes tipos de demencias (DCLewy, EA y DFT) y con deterioro subjetivo de memoria. La edad media fue mayor en el grupo de pacientes con DCLewy (76 ± 7 años) respecto al resto de grupos, encontrándose la menor edad media en el grupo de pacientes con DFT (64 ± 12 años). La puntuación en MMSE fue relativamente baja, siendo la menor en el grupo de pacientes con EA (13 ± 5 puntos) y la mayor en el grupo de pacientes con deterioro subjetivo de memoria (28 ± 2 puntos), y en los pacientes con DCLewy de 17 ± 6 puntos. Se encontró que la sensibilidad de la pérdida de este signo fue de 0,80 (IC95% 0,53 – 0,93) en pacientes con DCLewy respecto al resto de entidades descritas, y la especificidad fue de 0,76 (IC95% 0,57 – 0,88) en el grupo de DCLewy respecto al resto de pacientes sin DCLewy.

El estudio retrospectivo de Bruun et al. utilizó varias estructuras cerebrales para definir la sensibilidad y especificidad de la RM cerebral para la detección de diferentes tipos de demencia ⁽⁴¹⁾. Agrupó un total de 1213 pacientes con distintos tipos de demencia de dos cohortes prediseñadas. En primer lugar, valoró un índice combinando la corteza anterior frente a la posterior, hallando una sensibilidad de 0,59 y una especificidad de 0,95 para diferenciar la DFT de otro tipo de demencias. En segundo lugar, utilizó el índice de asimetría entre diferentes regiones cerebrales para diferenciar la afasia primaria progresiva (APP) de la variante conductual de la DFT, reportando una sensibilidad de 0,79 y una especificidad de 0,92. Por último, usó el polo temporal izquierdo para el diagnóstico de la APP de tipo semántico, con una sensibilidad

de 0,82 y una especificidad de 0,80. No se aportaban intervalos de confianza para los valores de sensibilidad y especificidad.

DISCUSIÓN

Esta revisión ha analizado la bibliografía existente a partir del año 2011 y hasta la actualidad sobre la precisión diagnóstica de la RM cerebral para el diagnóstico precoz de demencia en pacientes con sospecha de la misma o con DCL, así como su efectividad para diferenciar distintos tipos de demencia e identificar patrones de imagen característicos de esta patología.

Respecto a la efectividad de la RM cerebral en el diagnóstico de pacientes con EA, no se analizaron estudios que investigaran este aspecto, puesto que todos los estudios analizados incluían como muestra sujetos con DCL para investigar el papel de la RM cerebral en el diagnóstico precoz de demencia.

En cuanto al diagnóstico temprano de EA en pacientes diagnosticados de DCL, se encontraron diferentes resultados según la región cerebral analizada. En primer lugar, respecto a la medida del hipocampo, se encontró una sensibilidad media de 0,61 (IC95% 0,46 – 0,78) y una especificidad media de 0,69 (IC95% 0,55 – 0,81) en los estudios que utilizaban mediciones automatizadas, mientras que el único estudio que utilizó métodos manuales para la medición del volumen completo del hipocampo reportó una sensibilidad de 0,78 (IC95% 0,68 – 0,85) y una especificidad de 0,65 (IC95% 0,59 – 0,71). En global, la sensibilidad media encontrada en los diferentes estudios fue de 0,62 (IC95% 0,48 – 0,79) y la especificidad media de 0,70 (IC95% 0,55 – 0,80). En la revisión sistemática de Lombardi et al. de la que se extrajeron la mayoría de artículos analizados ⁽¹⁷⁾, se reportó una sensibilidad media de 0,73 y una especificidad media de 0,71, difiriendo levemente el valor de la sensibilidad respecto a lo encontrado en nuestro trabajo. Una de las posibles explicaciones es que dicha revisión sistemática incluía tres estudios más antiguos (de los años 1999, 2007 y 2008) que reportaban un valor de sensibilidad de 1, mientras que dichos estudios no fueron incluidos en nuestra revisión (el estudio más antiguo incluido en nuestro trabajo era del año 2013). Además, los estudios más antiguos pueden aportar valores de precisión diagnóstica diferentes por utilizar técnicas de imagen con menor resolución respecto a las técnicas de medida y los avances en neuroimagen que han tenido lugar los últimos años.

En cuanto al lóbulo temporal, en nuestro trabajo se encontró una sensibilidad media en los estudios de 0,65 (rango 0,45) y una especificidad de 0,69 (rango 0,32), usando métodos manuales de medición, y solo en los estudios de Pereira et al. y de Clerx et al. utilizando la escala MTA. En comparación, en otra revisión sistemática reciente centrada en la escala MTA para la medición de la atrofia del lóbulo temporal medial para diferenciar pacientes con EA de controles

sanos, se encontró una sensibilidad de 0,74 y una especificidad de 0,88⁽⁴²⁾. Por otro lado, en la revisión sistemática de Fink et al. del año 2020, la sensibilidad y especificidad reportadas para la atrofia del lóbulo temporal medial fueron también más elevadas, de 0,91 y 0,89 respectivamente⁽⁴³⁾. Estas diferencias las consideramos atribuibles en gran parte a la heterogeneidad en los métodos de medida del lóbulo temporal en los estudios analizados, y sobre todo porque en la revisión de Fink et al. se valoraba precisión diagnóstica de la RM para diferenciar pacientes con y sin demencia, sin tener en cuenta si presentaban diagnóstico previo de DCL. Igualmente, en la revisión previamente comentada de Lombardi et al.⁽¹⁷⁾, que incluyó siete estudios que analizaban el lóbulo temporal medial, se reportó una sensibilidad media de 0,64 y una especificidad media de 0,65 para el diagnóstico temprano de demencia en pacientes con DCL, siendo valores más similares a los encontrados en nuestra revisión dada la mayor semejanza con la población diana principal de nuestro trabajo.

Sobre el resto de estructuras cerebrales analizadas, no se han podido aportar datos suficientes para establecer conclusiones firmes dada la limitada cantidad de estudios. El único estudio incluido que valoraba la atrofia del lóbulo frontal aislada como marcador diagnóstico de la DFT mostraba una sensibilidad de 0,77 y una especificidad de 0,93, con una precisión total de 0,89. En comparación, un estudio prospectivo sobre la utilidad de las escalas visuales para la medición de la atrofia en distintas regiones cerebrales mostró utilidad para diferenciar pacientes con DFT variante conductual de los controles, con un área bajo la curva de 0,93⁽⁴⁴⁾ (si bien en este estudio no se valoraba únicamente la atrofia del lóbulo frontal, sino también otras áreas cerebrales). No obstante, respecto a la caracterización de la DFT, existen estudios previos que establecen una correlación entre el grado de atrofia de la región frontal y los síntomas conductuales, aunque dicha región no es la única afectada en este tipo de demencia⁽⁴⁵⁻⁴⁷⁾. En cuanto a la combinación de varias estructuras cerebrales para el diagnóstico, en otro artículo se estimó que el rendimiento diagnóstico de la RM cerebral para la predicción de demencia tipo EA en pacientes con DCL era moderado, con un área bajo la curva de 0,71⁽⁴⁸⁾, si bien no se aportaban datos concretos de sensibilidad y especificidad.

Los citados valores de sensibilidad y especificidad (0,67 y 0,68 respectivamente, según los resultados del metanálisis) no se consideran adecuados para utilizar la RM cerebral como prueba rutinaria para el diagnóstico precoz de demencia en pacientes con DCL, ya que deberían ser más elevados para considerar a dicha prueba como efectiva. Dichos valores contrastan con otras pruebas de imagen utilizadas recientemente como apoyo al diagnóstico de demencia, como el 18F-FDG PET, que muestra unos valores más aceptables de sensibilidad (0,76) y especificidad

(0,82) según una revisión sistemática de Cochrane ⁽⁴⁹⁾. Además, los valores de la razón de verosimilitud negativa (0,45, con IC95% 0,38 – 0,54) y positiva (2,0, con IC95% 1,69 – 2,42) no aportan una utilidad buena para considerar a la RM cerebral suficientemente discriminativa al comparar el cociente entre la tasa de falsos negativos y la de verdaderos negativos, y la tasa de verdaderos positivos y de falsos positivos, respectivamente.

En la guía de práctica clínica NICE más reciente sobre demencias (del año 2018) ⁽⁵⁰⁾, únicamente se contempla la realización de RM cerebral en caso de que el subtipo de demencia no esté claro y se sospeche demencia de tipo vascular. En el resto de casos, recomienda siempre una prueba de imagen estructural para descartar causas reversibles de deterioro cognitivo y ayudar en el diagnóstico del subtipo de demencia (sólo en los casos en los que no se pueda reconocer por la clínica del paciente), pero sin especificar si debe tratarse de una TC o de una RM cerebral.

Respecto a la capacidad de la RM cerebral para diferenciar distintos tipos de demencia, solo se encontró un estudio que analizara este aspecto, concluyendo que el índice de la región anterior frente a la posterior tendría mayor utilidad para diferenciar la DFT de otros tipos de demencia, el índice de asimetría para diferenciar la APP y DFT (en su variante conductual) y el índice del polo temporal izquierdo para detectar la variante semántica de la APP. Existe poca bibliografía existente sobre la efectividad de la prueba para diferenciar entre los distintos tipos de demencia, aunque sí es conocido que unas regiones cerebrales se afectan más que otras según el tipo de enfermedad ⁽⁵¹⁾. Dada la limitada cantidad de estudios que investiguen este punto, la amplia variabilidad de resultados, la escasa precisión con amplios rangos y la elevada heterogeneidad, no es posible extraer conclusiones relevantes.

Finalmente, es pertinente reseñar que este trabajo posee una serie de limitaciones. En primer lugar, destacamos como la mayor limitación el elevado índice de heterogeneidad entre los diferentes estudios, dada la utilización de diferentes metodologías para la medición de las estructuras cerebrales y los puntos de corte variables. Asimismo, el alto riesgo de sesgo en los estudios y el pequeño tamaño muestral de los mismos también se encuentran entre las limitaciones. Además, la ausencia de estudios suficientes no permite una extrapolación muy precisa de los resultados para cada estructura cerebral, con la excepción del hipocampo, donde sí se encontró un mayor número de resultados disponibles. Por otro lado, la falta de disponibilidad de suficientes datos para la realización de las tablas 2x2 en algunos estudios no permitió la agregación de todos los estudios analizados para el metanálisis, lo que también se interpreta como una limitación, sumado a la elevada heterogeneidad con valores de I^2 por encima del 60%, que no aporta gran fiabilidad a los resultados. Por otro lado, no se aportan

datos de variabilidad inter-observador o índice kappa en los estudios analizados, siendo un dato relevante en este tipo de pruebas de imagen, por lo que se considera una limitación añadida. Además, es precisa una formación de los profesionales en Radiología para la interpretación correcta de los resultados, por lo que es algo también a considerar en caso de utilizar esta prueba como diagnóstico precoz en pacientes con DCL o para intentar diferenciar entre distintos tipos de demencia en la práctica clínica habitual. A ello se suma la limitación del programa utilizado para la realización del metanálisis, que no pondera los estudios en función de su precisión, por lo que la medida agregada final podría estar distorsionada hacia aquellos estudios con mayor muestra.

Adicionalmente, como fortaleza del presente trabajo destaca la posibilidad que brinda para analizar con detalle la bibliografía más actualizada disponible sobre el papel de la RM cerebral en el diagnóstico de la demencia. En este sentido, aporta información sintetizada de utilidad para conocer qué estructuras cerebrales son las más relevantes para analizar y conclusiones sobre la precisión diagnóstica esperable de la RM cerebral. De esta forma, la aplicabilidad de esta revisión es aportar información para la elaboración de una guía de práctica clínica a nivel nacional, actualizando la última publicada del año 2011.

CONCLUSIONES

En conclusión, no existe hasta la fecha suficiente evidencia para establecer la RM cerebral como prueba de detección/diagnóstico de demencia en pacientes con DCL. Los valores de sensibilidad y especificidad obtenidos en el metanálisis (0,67 y 0,68 respectivamente) no alcanzan la suficiente precisión para poder discriminar a pacientes con demencia a través del uso de la RM cerebral. Además, la certeza de la evidencia según la herramienta GRADE se establece como muy baja. A ello se suman las limitaciones previamente mencionadas respecto a la elevada heterogeneidad entre los estudios analizados, que no permite extrapolar los valores de precisión diagnóstica de todos los estudios incluidos de forma rigurosa.

En un futuro, sería necesaria la realización de estudios adicionales con mayor homogeneidad en sus métodos, que permitan aportar mayor calidad en la comparación de sus resultados. En este sentido, sería de gran utilidad la creación de protocolos de imagen en nuestro país para conseguir una mayor comparabilidad entre pruebas diagnósticas interpretadas por diferentes profesionales y realizadas en centros distintos, tal y como se ha propuesto en un estudio en Canadá ⁽⁵²⁾. Igualmente, desarrollar investigaciones que incluyan la metodología de inteligencia artificial y aprendizaje automático podrían suponer aportaciones muy interesantes en este campo, así como el uso de técnicas de RM más novedosas, como se ha visto en un estudio reciente ⁽⁵³⁾. No obstante, es preciso reseñar que la utilidad de los hallazgos en neuroimagen para el diagnóstico siempre deben ir acompañados de una suficiente evidencia clínica, puesto que alteraciones aisladas en pruebas de neuroimagen sin acompañarse de una sospecha clínica no podrían ser valoradas como una confirmación diagnóstica por sí mismas.

BIBLIOGRAFÍA

1. Jack CR, Bennett DA, Blennow K, Carrillo MC, Dunn B, Haeberlein SB, et al. NIA-AA Research Framework: Toward a biological definition of Alzheimer's disease. *Alzheimers Dement* [Internet]. 2018 Apr 1 [cited 2022 May 22];14(4):535–62. Available from: <https://pubmed.ncbi.nlm.nih.gov/29653606/>
2. Scheltens P, Blennow K, Breteler MMB, de Strooper B, Frisoni GB, Salloway S, et al. Alzheimer's disease. *Lancet* [Internet]. 2016;388(10043):505–17. Available from: [http://dx.doi.org/10.1016/S0140-6736\(15\)01124-1](http://dx.doi.org/10.1016/S0140-6736(15)01124-1)
3. Tudor A, Vasile A, Trifu S, Cristea M. Morphological classification and changes in dementia (Review). *Exp Ther Med* [Internet]. 2021 Nov 9;23(1):33. Available from: <http://www.spandidos-publications.com/10.3892/etm.2021.10955>
4. Gauthier S, Rosa-Neto P, Morais J, Webster C. World Alzheimer Report 2021: Journey through the diagnosis of dementia. *World Alzheimer's Dis Int* [Internet]. 2021 [cited 2022 May 22];1–314. Available from: <https://www.alzint.org/u/World-Alzheimer-Report-2021.pdf>
5. VizHub - GBD Results [Internet]. [cited 2022 Jun 23]. Available from: <https://vizhub.healthdata.org/gbd-results/>).
6. Banerjee D, Muralidharan A, Hakim Mohammed AR, Malik BH. Neuroimaging in Dementia: A Brief Review. *Cureus* [Internet]. 2020 Jun 18;12(Mci):6–13. Available from: <https://www.cureus.com/articles/26629-neuroimaging-in-dementia-a-brief-review>
7. Dubois B, Hampel H, Feldman HH, Scheltens P, Aisen P, Andrieu S, et al. Preclinical Alzheimer's disease: Definition, natural history, and diagnostic criteria. *Alzheimer's Dement* [Internet]. 2016 Mar 1 [cited 2022 May 22];12(3):292–323. Available from: <https://onlinelibrary.wiley.com/doi/full/10.1016/j.jalz.2016.02.002>
8. Albert MS, Dekosky ST, Dickson D, Dubois B, Feldman HH, Fox NC, et al. The diagnosis of mild cognitive impairment 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):270–9.
9. Rittman T. Neurological update : neuroimaging in dementia. *J Neurol* [Internet]. 2020;267(11):3429–35. Available from: <https://doi.org/10.1007/s00415-020-10040-0>
10. Wahlund L. Psychiatry Research : Neuroimaging Structural brain imaging as a diagnostic tool in dementia , why and how ? *Psychiatry Res Neuroimaging* [Internet]. 2020;(September):111183. Available from: <https://doi.org/10.1016/j.pscychresns.2020.111183>
11. Young PNE, Estarellas M, Coomans E, Srikrishna M, Beaumont H, Maass A, et al. Imaging biomarkers in neurodegeneration: current and future practices. *Alzheimers Res Ther* [Internet]. 2020 Dec 27;12(1):49. Available from: <https://alzres.biomedcentral.com/articles/10.1186/s13195-020-00612-7>
12. Ausó E, Gómez-Vicente V, Esquiva G. Biomarkers for Alzheimer's Disease Early Diagnosis. *J Pers Med* [Internet]. 2020 Sep 4;10(3). Available from: <http://www.ncbi.nlm.nih.gov/pubmed/32899797>
13. Ciapponi A. QUADAS-2 : instrumento para la evaluación de la calidad de estudios de precisión diagnóstica. *Evidencia, Actual en la Práctica Ambulatoria* [Internet]. 2015;18(1):22–6. Available from: <https://doi.org/10.51987/evidencia.v18i1.6341>

14. Guyatt GH, Alonso-Coello P, Schünemann HJ, Djulbegovic B, Nothacker M, Lange S, et al. Guideline panels should seldom make good practice statements: guidance from the GRADE Working Group. *J Clin Epidemiol*. 2016;80(2016):3–7.
 15. Sanabria AJ, Rigau D, Rotaecche R, Selva A, Marzo-Castillejo M, Alonso-Coello P. Sistema GRADE: Metodología para la realización de recomendaciones para la práctica clínica. *Aten Primaria*. 2015;47(1):48–55.
 16. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*. 2021;372.
 17. Lombardi G, Crescioli G, Cavedo E, Lucenteforte E, Casazza G, Bellatorre A-G, et al. Structural magnetic resonance imaging for the early diagnosis of dementia due to Alzheimer’s disease in people with mild cognitive impairment. *Cochrane Database Syst Rev* [Internet]. 2020 Mar 2; Available from: <https://doi.wiley.com/10.1002/14651858.CD009628.pub2>
 18. Orso B, Mattei C, Arnaldi D, Massa F, Serafini G, Plantone D, et al. Clinical and MRI Predictors of Conversion From Mild Behavioural Impairment to Dementia. *Am J Geriatr Psychiatry*. 2020;28(7):755–63.
 19. Douaud G, Menke RAL, Gass A, Monsch AU, Rao A, Whitcher B, et al. Brain microstructure reveals early abnormalities more than two years prior to clinical progression from mild cognitive impairment to Alzheimer’s disease. *J Neurosci* [Internet]. 2013;33(5):2147–55. Available from: <https://pubmed.ncbi.nlm.nih.gov/23365250>
 20. Ferrari BL, De Carvalho Campos Neto G, Nucci MP, Mamani JB, Lacerda SS, Felício AC, et al. The accuracy of hippocampal volumetry and glucose metabolism for the diagnosis of patients with suspected Alzheimer’s disease, using automatic quantitative clinical tools. *Med (United States)*. 2019;98(45).
 21. Frölich L, Peters O, Lewczuk P, Gruber O, Teipel SJ, Gertz HJ, et al. Incremental value of biomarker combinations to predict progression of mild cognitive impairment to Alzheimer’s dementia. *Alzheimer’s Res Ther*. 2017;9(1):1–15.
 22. Monge Argilés JA, Blanco Cantó MA, Leiva Salinas C, Flors L, Muñoz Ruiz C, Sánchez Payá J, et al. Comparación de la capacidad diagnóstica precoz de los biomarcadores de la enfermedad de Alzheimer en resonancia magnética cerebral y líquido cefalorraquídeo. *Neurología*. 2014;29(7):397–401.
 23. Pereira JB, Cavallin L, Spulber G, Aguilar C, Mecocci P, Vellas B, et al. Influence of age, disease onset and ApoE4 on visual medial temporal lobe atrophy cut-offs. *J Intern Med*. 2014;275(3):317–30.
 24. Prestia A, Caroli A, Herholz K, Reiman E, Chen K, Jagust WJ, et al. Diagnostic accuracy of markers for prodromal Alzheimer’s disease in independent clinical series. *Alzheimers Dement* [Internet]. 2013 Nov;9(6):677–86. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/23375562>
 25. Moscoso A, Silva-Rodríguez J, Aldrey JM, Cortés J, Fernández-Ferreiro A, Gómez-Lado N, et al. Prediction of Alzheimer’s disease dementia with MRI beyond the short-term: Implications for the design of predictive models. *NeuroImage Clin*. 2019;23(April).
 26. Clerx L, van Rossum IA, Burns L, Knol DL, Scheltens P, Verhey F, et al. Measurements of medial temporal lobe atrophy for prediction of Alzheimer’s disease in subjects with
-

- mild cognitive impairment. *Neurobiol Aging* [Internet]. 2013;34(8):2003–13. Available from: <http://dx.doi.org/10.1016/j.neurobiolaging.2013.02.002>
27. Nesteruk M, Nesteruk T, Styczynska M, Mandecka M, Barczak A, Barcikowska M. Combined use of biochemical and volumetric biomarkers to assess the risk of conversion of mild cognitive impairment to Alzheimer's disease. *Folia Neuropathol*. 2016;54(4):369–74.
 28. Gaser C, Franke K, Klöppel S, Koutsouleris N, Sauer H. BrainAGE in Mild Cognitive Impaired Patients: Predicting the Conversion to Alzheimer's Disease. *PLoS One*. 2013;8(6).
 29. Silva Fuente-Alba C, Molina Villagra M. Likelihood ratio (razón de verosimilitud): definición y aplicación en Radiología. *Rev Argentina Radiol*. 2017;81(3):204–8.
 30. Vos SJ, van Rossum IA, Verhey F, Knol DL, Soininen H, Wahlund LO, et al. Prediction of Alzheimer disease in subjects with amnesic and nonamnesic MCI. *Neurology*. 2013/03/01. 2013;80(12):1124–32.
 31. Wood RA, Moodley KK, Lever C, Minati L, Chan D. Allocentric spatial memory testing predicts conversion from mild cognitive impairment to dementia: An initial proof-of-concept study. *Front Neurol*. 2016;7(DEC).
 32. Khan W, Westman E, Jones N, Wahlund LO, Mecocci P, Vellas B, et al. Automated Hippocampal Subfield Measures as Predictors of Conversion from Mild Cognitive Impairment to Alzheimer's Disease in Two Independent Cohorts. *Brain Topogr*. 2015;28(5):746–59.
 33. Eckerström C, Olsson E, Bjerke M, Malmgren H, Edman Å, Wallin A, et al. A combination of neuropsychological, neuroimaging, and cerebrospinal fluid markers predicts conversion from mild cognitive impairment to dementia. *J Alzheimer's Dis*. 2013;36(3):421–31.
 34. Ong KT, Villemagne VL, Bahar-Fuchs A, Lamb F, Langdon N, Catafau AM, et al. Aβ imaging with 18F-florbetaben in prodromal Alzheimer's disease: a prospective outcome study. *J Neurol Neurosurg Psychiatry*. 2014/06/28. 2015;86(4):431–6.
 35. Platero C, López ME, Carmen Tobar M del, Yus M, Maestu F. Discriminating Alzheimer's disease progression using a new hippocampal marker from T1-weighted MRI: The local surface roughness. *Hum Brain Mapp*. 2019;40(5):1666–76.
 36. Popescu SG, Whittington A, Gunn RN, Matthews PM, Glocker B, Sharp DJ, et al. Nonlinear biomarker interactions in conversion from mild cognitive impairment to Alzheimer's disease. *Hum Brain Mapp*. 2020;41(15):4406–18.
 37. Rhodius-Meester HFM, Koikkalainen J, Mattila J, Teunissen CE, Barkhof F, Lemstra AW, et al. Integrating Biomarkers for Underlying Alzheimer's Disease in Mild Cognitive Impairment in Daily Practice: Comparison of a Clinical Decision Support System with Individual Biomarkers. *J Alzheimer's Dis*. 2016;50(1):261–70.
 38. Ledig C, Schuh A, Guerrero R, Heckemann RA, Rueckert D. Structural brain imaging in Alzheimer's disease and mild cognitive impairment: biomarker analysis and shared morphometry database. *Sci Rep*. 2018;8(1):1–16.
 39. Prieto Del Val L, Cantero JL, Atienza M. Atrophy of amygdala and abnormal memory-related alpha oscillations over posterior cingulate predict conversion to Alzheimer's disease. *Sci Rep*. 2016;6(August):1–12.
-

40. Rizzo G, De Blasi R, Capozzo R, Tortelli R, Barulli MR, Liguori R, et al. Loss of Swallow Tail Sign on Susceptibility-Weighted Imaging in Dementia with Lewy Bodies. *J Alzheimer's Dis.* 2019;67(1):61–5.
41. Bruun M, Koikkalainen J, Rhodius-Meester HFM, Baroni M, Gjerum L, van Gils M, et al. Detecting frontotemporal dementia syndromes using MRI biomarkers. *NeuroImage Clin* [Internet]. 2019;22(February):101711. Available from: <https://doi.org/10.1016/j.nicl.2019.101711>
42. Park HY, Park CR, Suh CH, Shim WH, Kim SJ. Diagnostic performance of the medial temporal lobe atrophy scale in patients with Alzheimer's disease: a systematic review and meta-analysis. *Eur Radiol* [Internet]. 2021;31(12):9060–72. Available from: <https://doi.org/10.1007/s00330-021-08227-8>
43. Fink HA, Linskens EJ, Silverman PC, McCarten JR, Hemmy LS, Ouellette JM, et al. Accuracy of biomarker testing for neuropathologically defined Alzheimer disease in older adults with dementia: A systematic review. *Ann Intern Med.* 2020;172(10):669–77.
44. Illán-Gala I, Falgàs N, Friedberg A, Castro-Suárez S, Keret O, Rogers N, et al. Diagnostic Utility of Measuring Cerebral Atrophy in the Behavioral Variant of Frontotemporal Dementia and Association with Clinical Deterioration. *JAMA Netw Open.* 2021 Mar 11;4(3).
45. Cajanus A, Solje E, Koikkalainen J, Lötjönen J, Suhonen NM, Hallikainen I, et al. The Association Between Distinct Frontal Brain Volumes and Behavioral Symptoms in Mild Cognitive Impairment, Alzheimer's Disease, and Frontotemporal Dementia. *Front Neurol.* 2019;10(October):1–11.
46. Seeley WW, Crawford R, Rascofsky K, Kramer JH, Weiner M, Miller BL, et al. Frontal paralimbic network atrophy in very mild behavioral variant frontotemporal dementia. *Arch Neurol* [Internet]. 2008 Feb;65(2):249–55. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/18268196>
47. Convery R, Mead S, Rohrer JD. Review: Clinical, genetic and neuroimaging features of frontotemporal dementia. *Neuropathol Appl Neurobiol.* 2019;45(1):6–18.
48. Schmand B, Eikelenboom P, Van Gool WA. Value of diagnostic tests to predict conversion to alzheimer's disease in young and old patients with amnesic mild cognitive impairment. *J Alzheimer's Dis.* 2012;29(3):641–8.
49. Smailagic N, Vacante M, Hyde C, Martin S, Ukoumunne O, Sachpekidis C. 18 F-FDG PET for the early diagnosis of Alzheimer's disease dementia and other dementias in people with mild cognitive impairment (MCI). *Cochrane Database Syst Rev* [Internet]. 2015 Jan 28;(Mci). Available from: <https://doi.wiley.com/10.1002/14651858.CD010632.pub2>
50. National Institute for Health and Care Excellence. Dementia: assessment, management and support for people living with dementia and their carers. *NICE Guidel* [Internet]. 2018;(June 2018):2–43. Available from: www.nice.org.uk/guidance/ng97
51. Bhogal P, Mahoney C, Graeme-Baker S, Roy A, Shah S, Fraioli F, et al. The common dementias: A pictorial review. *Eur Radiol.* 2013;23(12):3405–17.
52. Duchesne S, Chouinard I, Potvin O, Fonov VS, Khademi A, Bartha R, et al. The Canadian Dementia Imaging Protocol: Harmonizing National Cohorts. *J Magn Reson Imaging.* 2019;49(2):456–65.

53. Inglese M, Patel N, Linton-Reid K, Loreto F, Win Z, Perry RJ, et al. A predictive model using the mesoscopic architecture of the living brain to detect Alzheimer’s disease. *Commun Med* [Internet]. 2022 Dec 20;2(1):70. Available from: <https://www.nature.com/articles/s43856-022-00133-4>

TABLAS Y FIGURAS

Tabla 1. Estrategia de búsqueda.

Estrategia de búsqueda en Medline	
1	magnetic resonance.mp. or Magnetic Resonance Spectroscopy/ (682286)
2	Magnetic Resonance Imaging.mp. or Magnetic Resonance Imaging/ (467722)
3	((Magnetic adj4 Resonance*) or (MR or NMR or MRI or FMRI)).tw. (653322)
4	1 or 2 or 3 (1371700)
5	DEMENTIA.mp. or DEMENTIA/ or FRONTOTEMPORAL DEMENTIA/ or DEMENTIA, VASCULAR/ or DEMENTIA, MULTI-INFARCT/ (113600)
6	ALZHEIMER DISEASE/ or alzheimer.mp. (95342)
7	lewy body disease.mp. or Lewy Body Disease/ (3573)
8	(Dementia or Amentia or Demention).tw. (95397)
9	((Vascular adj4 Dementia) or (Arteriosclerotic adj4 (Dementia or encephalopath* or leukoencephalopath*)) or (Binswanger adj3 (Disease or encephalopath*)) or CADASIL or (Subcortic* adj4 (encephalopath* or leukoencephalopath)) or ((multiinfarct* or multi-infarct) adj3 Dementia) or (Front* adj3 Lobe adj3 Dementia) or ((corticobasal or cortico basal or Front) adj4 (degenerati* PR dysfunction or Dementia)) or (FTD or FTLD)).tw. (13133)
10	5 or 6 or 7 or 8 or 9 (180154)
11	4 and 10 (22194)
12	limit 21 to (humans and yr="2011 -Current" and (meta analysis or "systematic review")) (345)
Estrategia de búsqueda en EMBASE	
	No.,Query,Results,Date
	#18,"#4 AND #16 AND [(cochrane review]/lim OR [systematic review]/lim OR [meta analysis]/lim)",866,30 Jun 2022
	#17,"#4 AND #16",53253,30 Jun 2022
	#16,"#5 OR #6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12 OR #13 OR #14 OR #15",466062,30 Jun 2022
	#15,"((cortico* OR front) NEAR/4 (degenerati* OR dysfunction OR dementia)):ti,ab,kw",3256,30 Jun 2022
	#14,"(front* NEAR/3 lobe NEAR/3 dementia):ti,ab,kw",531,30 Jun 2022
	#13,"((multiinfarct* OR 'multi infarct') NEAR/3 dementi*):ti,ab,kw",1239,30 Jun 2022
	#12,"(subcortic* NEAR/4 (encephalopath* OR leukoencephalopath*)):ti,ab,kw",1671,30 Jun 2022
	#11,"cadasil:ti,ab,kw",1758,30 Jun 2022
	#10,"(binswanger NEAR/3 (disease OR encephalopath*)):ti,ab,kw",519,30 Jun 2022
	#9,"(arteriosclerotic NEAR/4 (dementia OR encephalopath* OR leukoencephalopath*)):ti,ab,kw",275,30 Jun 2022
	#8,"(vascular NEAR/4 dementia):ti,ab,kw",10328,30 Jun 2022
	#7,"dementia:ti,ab,kw OR amentia:ti,ab,kw OR demention:ti,ab,kw",144856,30 Jun 2022
	#6,"'alzheimer disease' OR 'diffuse lewy body disease'",187434,30 Jun 2022
	#5,"'dementia'/exp OR 'dementia' OR 'frontotemporal dementia'/exp OR 'frontotemporal dementia' OR 'multiinfarct dementia'/exp",357270,30 Jun 2022
	#4,"#1 OR #2 OR #3",1511987,30 Jun 2022
	#3,"mr:ti,ab,kw OR nmr:ti,ab,kw OR mri:ti,ab,kw OR fmri:ti,ab,kw",927537,30 Jun 2022
	#2,"'nuclear magnetic resonance spectroscopy'/exp",115407,30 Jun 2022
	#1,"'nuclear magnetic resonance imaging'/exp",849471,30 Jun 2022

Tabla 2: Características principales de los estudios incluidos.

Autor	Título	Lugar de realización	Muestra (n)	Edad media (años)	% mujeres	MMSE basal	Diseño del estudio	Tiempo de seguimiento	Resumen de resultados	nº pacientes con evolución a demencia	Prueba de referencia
Bruun et al. (2019)	Detecting frontotemporal dementia syndromes using MRI biomarkers	Multicéntrico (Dinamarca, Finlandia, Italia, Holanda, Reino Unido)	1213: 614 del Amsterdam Dementia Cohort (ADC) y 599 del PredictND (cohorte multicéntrica)	EA 68±8, DCLewy 69±8, Dva 71±8, Otra 73±9, DCL 67±8, DCS 62±9, DFT 64±7.	EA 56%, DCLewy 15%, Dva 38%, Otra 51%, DCL 38%, DCS (det cogn subj) 55%, DFT 43%.	EA 22±5, DCLewy 24±4, Dva 24±4, Otra 25±4, DCL 27±2, DCS 29±1, DFT 24±5.	De cohortes retrospectivo	12 meses (mínimo)	Índice región anterior VS posterior: para diferenciar la DFT de otros tipos de demencia. Índice de asimetría: para separar APP y DFTco. Índice del polo temporal izquierdo: detección de APPse.	-	NIA-AA para EA; criterios de Neary & Snowden et al. o de Mckhann et al. (para DFT); de Rascovsky et al. (para DFT variante conductual), de Gorno-Tempini (para APP variante semántica y no fluente), de NINDS-AIREN (para demencia vascular) y de McKeith (para DCL)
Clerx et al. (2013)	Measurements of medial temporal lobe atrophy for prediction of Alzheimer's disease in subjects with mild cognitive impairment	Holanda	328 (156 de una cohorte multicéntrica y 172 de un centro médico)	70,6±7,6	51,50%	27,0±2,5	De cohortes prospectivo	2 años	AUC fue más alta para la medición automatizada del hipocampo (0,71) y la medición manual (0,71), y más baja para la puntuación MTA (0,65) y el ventrículo lateral (0,60). Las medidas volumétricas del hipocampo son los mejores predictores de conversión de EA en sujetos con DCL.	91 (27,7%) evolucionaron a EA, y 20 (6%) a otro tipo de demencias (4 DFT, 6 DCLewy, 1 Dva, 1 otro tipo)	Criterios DSM-IV y NINCDS/ADRDA
Douaud et al. (2013)	Brain microstructure reveals early abnormalities more than two years prior to clinical progression from mild cognitive impairment to Alzheimer's Disease	Reino Unido y Suiza	56 pacientes con DCL (13 con DCL progresivo, 22 con DCL estable y 21 que no cumplían dichos criterios).	DCL estable: 69±9 DCL progresivo: 76±6	DCL estable: 50% DCL progresivo: 23%	DCL estable: 28.5±1.5 DCL progresivo: 28±1.5	De cohortes prospectivo	-	El volumen de sustancia gris se mostró más pequeño en pacientes con DCL progresivo respecto a DCL estable, aunque no presentaba una sensibilidad considerable para la predicción del desarrollo de demencia en pacientes con DCL (39%).	13 de los 56 pacientes con DCL progresaron a EA (23,2%).	Criterios DSM-IV y NINCDS/ADRDA

Eckerström et. al. (2013)	A combination of neuropsychological, neuroimaging, and cerebrospinal fluid markers predicts conversion from mild cognitive impairment to dementia	Suecia	42 sujetos seleccionados de una cohorte de otro estudio (Gothenburg MCI study), 21 con DCL estable y 21 con DCL con evolución a demencia + 26 controles	DCL conversores a demencia no-EA: 69,3±6,3 DCL conversores a EA: 70,0±6,5 DCL con Dx final de DCL: 66,4±6.8 Controles: 68,0±5,4	DCL conversores a demencia no-EA: 71,4% DCL conversores a EA: 69,2% DCL con Dx final de DCL: 42,8% Controles: 69,2%	DCL conversores a demencia no-EA: 27,2±2,0 DCL conversores a EA: 28,0±1.1 DCL con Dx final de DCL: 28,3±1.8 Controles: 29,4±0,9	De cohortes prospectivo	2 años	En los pacientes con DCL que convierten a cualquier tipo de demencia, el volumen del hipocampo izquierdo tiene mayor sensibilidad que el del derecho (0,86 VS 0,81), ocurriendo lo mismo con la E (0,66 VS 0,57). Sin embargo, la E en el grupo de pacientes que convierten a EA es mayor, sobre todo en el hipocampo derecho respecto al izquierdo (0,81 VS 0,75).	21 de los 42 sujetos convierten a demencia, de los cuales 13 a EA (30,9%), y 19% a otras demencias (4 a Dva, 2 a tipo mixto EA/Dva y 2 a DFT).	Criterios NINCDS/ADRDA
Ferrari et. al. (2019)	The accuracy of hippocampal volumetry and glucose metabolism for the diagnosis of patients with suspected Alzheimer's disease, using automatic quantitative clinical tools	Brasil	38 pacientes con sospecha de AD	68.3±11.2	57.9%	-	De cohortes retrospectivo	23±11 meses (media)	Para el diagnóstico de EA y DCL, el volumen del hipocampo tiene baja S (13%) y alta E (95%) si se considera un punto de corte < p5, lo que llevaría a muchos FN. Con un punto de corte de 29,03 (derivado de la curva ROC), la S mejora al 63%, manteniendo una E aceptable del 68%, con una precisión diagnóstica del 66% (comparada con el 61% usando como punto de corte el p5).	16 pacientes desarrollaron EA o presentaron DCL de alto riesgo (42,1%).	Criterios NIA-AA para EA.
Frölich et. al. (2017)	Incremental value of biomarker combinations to predict progression of mild cognitive impairment to Alzheimer's dementia	Alemania	115 pacientes con DCL de un estudio de cohortes multicéntrico ("Dementia Competence Network").	65.7±9,03	41,70%	27,0±2.12	De cohortes prospectivo	Hasta 3 años, con una media de seguimiento de 2.1 años	El volumen medio de hipocampo derecho e izquierdo puede predecir el desarrollo de demencia tipo EA con una S del 63% y una E del 82%	28 progresaron a demencia tipo EA (24,3%)	Ocurrencia de nuevos déficits en actividades básicas o instrumentales de la vida diaria (BADL score > 4).
Gaser et. al. (2013)	BrainAGE in Mild Cognitive Impaired Patients: Predicting the Conversion to Alzheimer's Disease	Alemania	195 pacientes con DCL de la base de datos ADNI	DCL progresivo temprano: 73,9 DCL progresivo tardío: 75,2 DCL estable: 76,4	DCL progresivo temprano: 43,1% DCL progresivo tardío: 36% DCL estable: 20,9%	DCL progresivo temprano: 26,5 DCL progresivo tardío: 26,8 DCL estable: 27,7	De cohortes prospectivo	3 años	El volumen basal del hipocampo en pacientes con DCL puede predecir el desarrollo de EA, con una S mayor en el lado derecho respecto al izquierdo (0,84 VS 0,52) y una E mayor en el izquierdo que en el derecho (0,81 VS 0,42).	133 pacientes (68,2%)	Criterios NINCDS/ADRDA

Khan et. al. (2015)	Automated hippocampal subfield measures as predictors of conversion from mild cognitive impairment to Alzheimer's Disease in two independent cohorts	Multicéntrico: Reino Unido, Suecia, Italia, Francia, Grecia, Polonia, Finlandia	447 pacientes con DCL, 291 EA y 331 controles de la base de datos ADNI y del estudio AddNeuroMed (de la UE).	EA: 75.4 ± 7.0 DCL que convierten en EA: 74.1 ± 6.6 DCL estable: 75.1 ± 7.0 Controles: 75.0 ± 5.7	EA: 55% DCL que convierten en EA: 40% DCL estable: 39,5% Controles: 49,8%	EA: 22.4 ± 3,4 DCL que convierten en EA: 26,5 ± 1,8 DCL estable: 27.1 ± 1,7 Controles: 29,1 ± 1,1	De cohortes prospectivo	1 año	La medición del volumen de hipocampo puede predecir el desarrollo de EA en pacientes con DCL en un seguimiento de 1 año, siendo algo más precisa la medición automatizada de distintas subregiones del hipocampo que el volumen total.	90 pacientes de los que presentaban DCL (447) evolucionaron a EA (20%)	Criterios DSM-IV y NINCDS/ADRDA
Ledig et. al. (2018)	Structural brain imaging in Alzheimer's disease and mild cognitive impairment: biomarker analysis and shared morphometry database	Reino Unido, Suecia	5074 imágenes de RM de la base de datos ADNI (1674 basales, 3400 de seguimiento). Pacientes con DCL: 343	Mediana (mín; máx): 74,6 (48.1; 91,4)	45,70%	Mediana (mín; máx): 27 (18; 30)	De cohortes prospectivo	2 años	La región cerebral más sensible para diferenciar DCL progresivo de estable fue la amígdala (S 63%, E 68%), y la más específica fue el ventrículo lateral (E 74%, S 54%)	177 presentan DCL progresivo de los 343 pacientes con DCL basal (52%)	Criterios NINCDS/ADRDA
Monge et. al. (2013)	Comparación de la capacidad diagnóstica precoz de los biomarcadores de la enfermedad de Alzheimer en resonancia magnética cerebral y líquido cefalorraquídeo	España	30 pacientes con DCL	DCL: 73,25 ± 6,63 EA: 72,56 ± 7,12	DCL: 53,3% EA: 66,7%	DCL: 24 ± 2,4 EA: 23 ± 1,2	De cohortes prospectivo	2 años	La medición de la atrofia del lóbulo temporal medial en RM tiene una S del 67% y una E del 53% para el diagnóstico precoz de EA en pacientes con DCL, aunque la precisión es mayor para los biomarcadores en LCR. La combinación de ambas técnicas consigue una precisión diagnóstica muy alta, tanto para diagnosticar como para excluir precozmente EA en pacientes con DCL.	15 pacientes (50%) progresaron a EA, el resto se mantuvieron con diagnóstico de DCL	Criterios NINCDS/ADRDA

Moscoso et. al. (2019)	Prediction of Alzheimer's disease dementia with MRI beyond the short-term: Implications for the design of predictive models	España	230 con EA y 230 controles de la base de datos ADNI, pero el seguimiento a 5 años solo lo completan 248 pacientes con DCL	ADNI: EA: 75 [55-90] DCL estable (2 años seg): 75 [55-88] DCL progresivo (2 años seg): 75 [56-88] ADNIGO/2: EA: 75 [56-90] DCL estable (2 años seg): 71 [55-85] DCL progresivo (2 años seg): 73 [57-85]	ADNI: EA: 47,6% DCL estable (2 años seg): 36,6% DCL progresivo (2 años seg): 42,7% ADNIGO/2: EA: 46,2% DCL estable (2 años seg): 46,9% DCL progresivo (2 años seg): 45,5%	ADNI: EA: 23 [20-27] DCL estable (2 años seg): 28 [24-30] DCL progresivo (2 años seg): 26 [23-30] ADNIGO/2: EA: 23 [19-26] DCL estable (2 años seg): 28 [24-30] DCL progresivo (2 años seg): 26 [24-30]	De cohortes prospectivo	3-5 años	La RM es altamente predictiva de EA a medio plazo (2 años) en el DCL progresivo, aumentando la especificidad y el poder de discriminación a los 5 años de seguimiento.	-	Criterios NINCDS/ADRDA para EA probable
Nesteruk et. al. (2016)	Combined use of biochemical and volumetric biomarkers to assess the risk of conversion of mild cognitive impairment to Alzheimer's disease	Polonia	40 pacientes con DCL	63,17 ± 9,56	55%	27,50 ± 1,73	De cohortes prospectivo	2 años	Los valores volumétricos en RM tienen una S del 88,9% y una E del 90,3% para la predicción de EA en pacientes con DCL. Las áreas que mostraron diferencia estadísticamente significativa en convertidores VS no convertidores fueron el hipocampo derecho y el izquierdo, el córtex entorrinal izquierdo y el giro temporal medio derecho. La combinación con marcadores de LCR no aumenta la sensibilidad, pero sí la especificidad.	9 pacientes con DCL (22,5%) convierten a EA	Criterios NIA-AA para EA
Ong et. al. (2015)	Aβ imaging with 18F-florbetaben in prodromal Alzheimer's disease: a prospective outcome study	Australia	45 pacientes con DCL	71,1 ± 6,1	-	27.6±1.8	De cohortes prospectivo	4 años	El volumen del hipocampo, medido con RM, muestra una S del 50% y una E del 64% para la predicción de EA en pacientes con DCL. Además, según avanza la neurodegeneración, la atrofia del hipocampo parece correlacionarse con el deterioro de la memoria.	20 pacientes progresan a EA (44,4%) tras los dos primeros años de seguimiento.	Criterios NINCDS/ADRDA

Orso et. al. (2019)	Clinical and MRI predictors of conversion from mild behavioural impairment to dementia	Italia	113 sujetos con síntomas conductuales	Grupo A: 65,4±7,9 Grupo B: 66,6±6,4	Grupo A: 26,8% Grupo B: 29,8%	Grupo A: 28,4±2,3 Grupo B: 28,0±1,4	De cohortes retrospectivo	4 años	En pacientes con alteraciones de conducta leves, un seguimiento retrospectivo a 4 años mostró para la RM una sensibilidad de 76,9% y una especificidad de 93,1% para el diagnóstico de DFT variante conductual.	Grupo A (n=56): 14 DFT y 4 EA (no analizada S y E en este subgrupo) Grupo B (n=57): 13 DFT, ninguno EA	No constan criterios diagnósticos
Pereira et. al. (2014)	Influence of age, disease onset and ApoE4 on visual medial temporal lobe atrophy cut-offs	Multicéntrico: Suecia, Italia, Francia, Grecia, Polonia, Finlandia, Reino Unido	1147 sujetos de las cohortes ADNI y AddNeuroMed (345 controles, 480 con DCL y 322 con EA).	EA: 75,6 ± 7,1 DCL progresivo: 74,3 ± 6,5 DCL estable: 75,1 ± 6,9	EA: 54,6% DCL progresivo: 41% DCL estable: 39,5%	EA: 21,5 ± 5,4 DCL progresivo: 26,5 ± 1,8 DCL estable: 27,1 ± 1,7	De cohortes prospectivo	1 año	Existen variables clínicas, demográficas y genéticas que pueden influir en la clasificación de las puntuaciones de corte de la escala MTA para la medición del hipocampo, lo que podría conducir a veces a un diagnóstico erróneo. El punto de corte en la escala MTA ≥1,5 es más sensible (84,5%) y el punto de corte según la edad es más específico (83,2%).	95 sujetos de los 480 que presentaban DCL progresaron a EA (19,8%) en 1 año de seguimiento	Criterios NINCDS/ADRDA
Platero et. al. (2019)	Discriminating Alzheimer's disease progression using a new hippocampal marker from T1-weighted MRI: The local surface roughness	España	97 pacientes con DCL	DCL estable: 73,2 ± 5,2 DCL progresivo: 75,6 ± 4,9	DCL estable: 64% DCL progresivo: 61%	DCL estable: 27,0 ± 2,3 DCL progresivo: 25,8 ± 3,1	De cohortes prospectivo	3 años	La rugosidad superficial del hipocampo fue el marcador de RM que más sensibilidad (77,4%) y especificidad (83,2%) que la medición del volumen de hipocampo normalizado o la medición de la rugosidad superficial.	36 de los 97 sujetos con DCL progresaron a EA (37,1%)	Criterios NIA-AA para EA
Popescu et. al. (2020)	Nonlinear biomarker interactions in conversion from mild cognitive impairment to Alzheimer's disease	Reino Unido	206 sujetos con DCL (158 estable y 48 progresivo), de la base de datos ADNI	DCL estable: 71,32±7,24 DCL progresivo: 73,35±6,76	DCL estable: 41,77% DCL progresivo: 45,83%	-	De cohortes prospectivo	3 años (mínimo)	El volumen de hipocampo de forma independiente puede predecir la evolución a EA en sujetos con DCL con una S del 66.6% y una E del 66,6%. El modelo con mayor rendimiento incluía una interacción entre amiloide-β-PET y P-tau, teniendo en cuenta el volumen del hipocampo (S 77%, E 76,2%).	48 pacientes con DCL de los 206 analizados presentaban DCL progresivo (23,3%).	Criterios NINCDS/ADRDA para EA probable

Prestia et. al. (2013)	Diagnostic accuracy of markers for prodromal Alzheimer's disease in independent clinical series	Multicéntrico: Italia, Reino Unido, EEUU	103 pacientes con DCL de la base de datos ADNI y de una clínica italiana (TOMC). 52 con DCL progresivo y 51 con DCL estable.	ADNI: DCL estable: 75 ± 8 DCL progresivo: 75 ± 8 TOMC: DCL estable: 72 ± 8 DCL progresivo: 71 ± 8	ADNI: DCL estable: 39% DCL progresivo: 42% TOMC: DCL estable: 50% DCL progresivo: 67%	ADNI: DCL estable: 27 ± 2 DCL progresivo: 28 ± 2 TOMC: DCL estable: 26 ± 2 DCL progresivo: 26 ± 2	De cohortes prospectivo	2-3 años	La sensibilidad del volumen de hipocampo en RM para predecir el desarrollo de EA en pacientes con DCL en las cohortes ADNI y TOMC fue 46% y 28% respectivamente, y la especificidad de 76% y 94%.	42 de los 93 sujetos con DCL que se incluyeron finalmente progresaron a EA (45%). 42% en la cohorte de ADNI y 50% en la cohorte de TOMC.	Criterios NINCDS/ADRDA
Prieto et. al. (2016)	Atrophy of amygdala and abnormal memory-related alpha oscillations over posterior cingulate predict conversion to Alzheimer's disease	España	60 sujetos: 26 controles y 34 con DCL amnésico (16 progresaron a EA y 18 DCL estable).	DCL estable: 68,4 ± 7,1 DCL progresivo: 69,7 ± 6,5	DCL estable: 72,2% DCL progresivo: 56,2%	DCL estable: 26,8 ± 2,1 DCL progresivo: 26,4 ± 2,7	De cohortes prospectivo	2 años	Los pacientes con DCL que progresaron a demencia tipo EA presentaron mayor grado de atrofia en la amígdala derecha, con una sensibilidad de 74,4% y una especificidad de 79,2%. Se obtuvo mayor precisión diagnóstica al combinar el volumen de amígdala con marcadores en EEG (desincronización en ritmo alfa sobre la corteza cingulada posterior), con una sensibilidad del 78% y una especificidad del 82%.	16 de los 34 pacientes con DCL progresaron a EA (47%)	No específica
Rhodus et. al. (2016)	Integrating biomarkers for underlying Alzheimer's Disease in mild cognitive impairment in daily practice: comparison of a clinical decision support system with individual biomarkers	Holanda, Finlandia	211 pacientes con DCL	DCL estable: 68 ± 6 DCL progresivo a EA: 72 ± 7	DCL estable: 31% DCL progresivo a EA: 55%	DCL estable: 27 ± 2 DCL progresivo a EA: 26 ± 3	De cohortes prospectivo	3 años (mediana)	El uso de la herramienta PredictAD para el diagnóstico de EA según la RM cerebral presentó resultados comparables en cuanto a precisión diagnóstica que el resto de biomarcadores individuales (puntuación MTA para medición de la atrofia del lóbulo temporal medial) y los criterios diagnósticos NIA-AA.	139 pacientes progresaron a demencia tipo EA (66%)	Criterios NIA-AA para EA

Rizzo et. al. (2019)	Loss of swallow tail sign on susceptibility-weighted imaging in Dementia with Lewy Bodies	Italia	44 pacientes (15 con DCLewy, 11 con EA, 8 con DFT y 10 con deterioro subjetivo de memoria)	DCLewy: 76±7 EA: 74±8 DFT: 64±12 Det.subjetivo de memoria: 67±9	DCLewy: 46,7% EA: 54,5% DFT: 62,5% Det.subjetivo de memoria: 50%	DCLewy: 17±6 EA: 13±5 DFT: 19±6 Det.subjetivo de memoria: 28±2	De cohortes prospectivo	1 año (mínimo)	Pérdida del signo de la cola de golondrina uni o bilateral: en 80% de pacientes con DCLewy, 36% de pacientes con EA, 25% de pacientes con DFT y 10% de pacientes con det.subjetivo de memoria	-	Criterios NIA-AA para EA; criterios de Neary & Snowden et al. para DFT; criterios de McKeith para DCLewy.
Vos et. al. (2013)	Prediction of Alzheimer disease in subjects with amnesic and nonamnesic mild cognitive impairment	Multicéntrico (Países Bajos, Holanda, Finlandia, Suecia, Alemania, Grecia, Italia, Reino Unido)	625 pacientes con DCL (399 amnésico y 226 no amnésico), pero solo se mide el volumen de hipocampo en 388 sujetos. De la base de datos ADNI.	DCL no amnésico: 70,7±7,6 DCL amnésico: 70,7±7,8	DCL no amnésico: 49% DCL amnésico: 56%	DCL no amnésico: 27,5±2,1 DCL amnésico: 26,5±2,5	De cohortes prospectivo	1-5 años	La RM puede ser de utilidad para la predicción de EA en sujetos con DCL. La sensibilidad en sujetos con DCL amnésico es mayor y la especificidad es menor respecto a sujetos con DCL no amnésico.	La progresión a demencia tipo EA fue mayor en el grupo de DCL amnésico (38%) que DCL no amnésico (20%).	Criterios de EA según DSM-IV y los criterios NINCDS/ADRDA
Wood et. al. (2016)	Alloentric spatial memory testing predicts conversion from mild cognitive impairment to dementia: an initial proof-of-concept study	Reino Unido, Italia	15 pacientes con DCL	DCL estable: 65,2 ± 3,4 DCL progresivo a EA: 71,7 ± 3,0	DCL estable: 33,3% DCL progresivo a EA: 22,2%	DCL estable: 27,33 ± 0,21 DCL progresivo a EA: 27,89 ± 0,42	De cohortes prospectivo	2 años	La medición del volumen del hipocampo, en relación con el volumen intracraneal total, mostró una sensibilidad del 85,7% y una especificidad del 66,7% para la predicción del desarrollo de demencia tipo EA en pacientes con DCL.	9 de 15 pacientes progresaron a EA (60%)	Criterios NIA-AA para EA

Tabla 3: Valores de sensibilidad y especificidad de los estudios incluidos.

Autor	Título	Sensibilidad	Especificidad
Bruun et. al. (2019)	Detecting frontotemporal dementia syndromes using MRI biomarkers	Índice región anterior VS posterior: 0,59 (para diferenciar DFT de otras demencias no DFT) Índice de asimetría: 0,79 (para separar la APP de la DFT variante conductual) Índice del polo temporal izquierdo: 0,82 (para detectar la APP semántica)	Índice región anterior VS posterior: 0,95 Índice de asimetría: 0,92 Índice del polo temporal izquierdo: 0,80
Clerx et. al. (2013)	Measurements of medial temporal lobe atrophy for prediction of Alzheimer's disease in subjects with mild cognitive impairment	Hipocampo (manual): 0,78 (0,68-0,85) Hipocampo (automatizada): 0,66 Lóbulo temporal medial (escala visual MTA-score): 0,66 Ventrículo lateral: 0,53	Hipocampo: 0,65 (0,59-0,71) Hipocampo: 0,77 Lóbulo temporal medial: 0,64 Ventrículo lateral: 0,68
Douaud et. al. (2013)	Brain microstructure reveals early abnormalities more than two years prior to clinical progression from mild cognitive impairment to Alzheimer's Disease	Volumen de sustancia gris: 0,39	Volumen de sustancia gris: 0,82
Eckerström et. al. (2013)	A combination of neuropsychological, neuroimaging, and cerebrospinal fluid markers predicts conversion from mild cognitive impairment to dementia	DCL que convierten a cualquier demencia (medición manual): Hipocampo derecho: 0,81 Hipocampo izquierdo: 0,86 DCL que convierten a EA (medición manual): Hipocampo derecho: 0,69 Hipocampo izquierdo: 0,85	DCL que convierten a cualquier demencia: Hipocampo derecho: 0,57 Hipocampo izquierdo: 0,66 DCL que convierten a EA: Hipocampo derecho: 0,81 Hipocampo izquierdo: 0,75
Ferrari et. al. (2019)	The accuracy of hippocampal volumetry and glucose metabolism for the diagnosis of patients with suspected Alzheimer's disease, using automatic quantitative clinical tools	Medición automatizada del volumen de hipocampo: Para punto de corte percentil 29,03: 0,63 (0,38-0,82) Para percentil<5: 0,13	Medición automatizada del volumen de hipocampo: Para corte de p29,03: 0,68 (0,47-0,84) Para percentil<5: 0,95
Frölich et. al. (2017)	Incremental value of biomarker combinations to predict progression of mild cognitive impairment to Alzheimer's dementia	Volumen de hipocampo (media del derecho y el izquierdo, medición automatizada): 0,63 (0,45-0,80)	Volumen de hipocampo: 0,82 (0,72-0,88)
Gaser et. al. (2013)	BrainAGE in Mild Cognitive Impaired Patients: Predicting the Conversion to Alzheimer's Disease	Medición automatizada: Volumen de hipocampo izquierdo: 0,52 Volumen de hipocampo derecho: 0,84	Medición automatizada: Volumen de hipocampo izquierdo: 0,81 Volumen de hipocampo derecho: 0,42
Khan et. al. (2015)	Automated hippocampal subfield measures as predictors of conversion from mild cognitive impairment to Alzheimer's Disease in two independent cohorts	Volumen total de hipocampo: 0,77 (0,67-0,84) (medición automatizada)	Volumen total de hipocampo: 0,50 (0,45-0,55)
Ledig et. al. (2018)	Structural brain imaging in Alzheimer's disease and mild cognitive impairment: biomarker analysis and shared morphometry database	Para diferenciar pacientes con DCL progresivo de estable: 0,72 (todas las estructuras cerebrales evaluadas en conjunto)	Para diferenciar pacientes con DCL progresivo de estable: 0,64
Monge et. al. (2013)	Comparación de la capacidad diagnóstica precoz de los biomarcadores de la enfermedad de Alzheimer en resonancia magnética cerebral y líquido cefalorraquídeo	Atrofia temporal medial (medición visual con el método de Korf et al.): 0,67	Atrofia temporal medial: 0,53

Moscoso et. al. (2019)	Prediction of Alzheimer's disease dementia with MRI beyond the short-term: Implications for the design of predictive models	Hipocampo: 0,83 (0,76-0,89) (2 años), 0,79 (0,72-0,84) (5 años) Córtex entorrinal: 0,77 (0,69-0,84) (2 años), 0,78 (0,71-0,84) (5 años) Modelo RM: 0,86 (0,79-0,81) (2 años), 0,78 (0,71-0,84) (5 años)	Hipocampo: 0,52 (0,43-0,61) (2 años), 0,63 (0,51-0,74) (5 años) Córtex entorrinal: 0,50 (0,39-0,58) (2 años), 0,66 (0,54-0,77) (5 años) Modelo RM: 0,56 (0,47-0,66) (2 años), 0,71 (0,60-0,81) (5 años)
Nesteruk et. al. (2016)	Combined use of biochemical and volumetric biomarkers to assess the risk of conversion of mild cognitive impairment to Alzheimer's disease	Combinación del volumen de varias estructuras cerebrales: 0,889	Combinación del volumen de varias estructuras cerebrales: 0,903
Ong et. al. (2015)	Aβ imaging with 18F-florbetaben in prodromal Alzheimer's disease: a prospective outcome study	Medición automatizada del volumen del hipocampo: 0,50 (0,29-0,71)	Medición automatizada del volumen del hipocampo: 0,64 (0,44-0,80)
Orso et. al. (2019)	Clinical and MRI predictors of conversion from mild behavioural impairment to dementia	Atrofia frontal aislada: 0,769	Atrofia frontal aislada: 0,931
Pereira et. al. (2014)	Influence of age, disease onset and ApoE4 on visual medial temporal lobe atrophy cut-offs	Para diagnóstico de EA: Para punto de corte MTA \geq 1,5: 0,85 (0,80-0,88) Para punto de corte por edad: 0,70 (0,65-0,75)	Para diagnóstico de EA: Para punto de corte MTA \geq 1,5: 0,68 (0,63-0,73) Para punto de corte por edad: 0,83 (0,79-0,87)
Platero et. al. (2019)	Discriminating Alzheimer's disease progression using a new hippocampal marker from T1-weighted MRI: The local surface roughness	Para diferenciar DCL estable VS DCL progresivo (mediciones automatizadas): Volumen de hipocampo normalizado: 0,68 (0,53-0,82) Rugosidad superficial: 0,675 Rugosidad superficial local: 0,774	Para diferenciar DCL estable VS DCL progresivo: Volumen de hipocampo normalizado: 0,68 (0,56-0,79) Rugosidad superficial: 0,685 Rugosidad superficial local: 0,723
Popescu et. al. (2020)	Nonlinear biomarker interactions in conversion from mild cognitive impairment to Alzheimer's disease	Volumen de hipocampo (automatizado): 0,666.	Volumen de hipocampo (automatizado): 0,666.
Prestia et. al. (2013)	Diagnostic accuracy of markers for prodromal Alzheimer's disease in independent clinical series	ADNI: Volumen de hipocampo semi-automatizado: 0,46 Volumen de hipocampo automatizado: 0,50 TOMC: Volumen de hipocampo semi-automatizado: 0,28 Volumen de hipocampo automatizado: 0,44	ADNI: Volumen de hipocampo semi-automatizado: 0,76 Volumen de hipocampo automatizado: 0,61 TOMC: Volumen de hipocampo semi-automatizado: 0,94 Volumen de hipocampo automatizado: 0,73
Prieto et. al. (2016)	Atrophy of amygdala and abnormal memory-related alpha oscillations over posterior cingulate predict conversion to Alzheimer's disease	Volumen de amígdala derecha: 0,744	Volumen de amígdala derecha: 0,792
Rhodus et. al. (2016)	Integrating biomarkers for underlying Alzheimer's Disease in mild cognitive impairment in daily practice: comparison of a clinical decision support system with individual biomarkers	Atrofia del lóbulo temporal medial (manual): 0,40 Herramienta PredictAD (DSI \geq 0,5): 0,67 (0,66-0,68)	Atrofia del lóbulo temporal medial (manual): 0,85 Herramienta PredictAD (DSI \geq 0,5): 0,62 (0,61-0,64)
Rizzo et. al. (2019)	Loss of swallow tail sign on susceptibility-weighted imaging in Dementia with Lewy Bodies	Pérdida del signo uni o bilateral: DCLewy VS EA: 0,80 DCLewy VS DFT: 0,80 DCLewy VS det.memoria subjetivo: 0,80 DCLewy VS no-DCLewy: 0,80 (0,53-0,93)	Pérdida del signo uni o bilateral: DCLewy VS EA: 0,64 DCLewy VS DFT: 0,75 DCLewy VS det.memoria subjetivo: 0,90 DCLewy VS no-DCLewy: 0,76 (0,57-0,88)
Vos et. al. (2013)	Prediction of Alzheimer disease in subjects with amnesic and nonamnesic mild cognitive impairment	Poder predictivo tras 2 años de seguimiento, midiendo volumen de hipocampo (automatizado): DCL no amnésico: 0,39 (0,16-0,61) DCL amnésico: 0,63 (0,51-0,74)	Poder predictivo tras 2 años de seguimiento, midiendo volumen de hipocampo (automatizado): DCL no amnésico: 0,88 (0,82-0,95) DCL amnésico: 0,72 (0,64-0,80)
Wood et. al. (2016)	Allocentric spatial memory testing predicts conversion from mild cognitive impairment to dementia: an initial proof-of-concept study	Volumen total de hipocampo (automatizado): 0,86 (0,50-0,99)	Volumen total de hipocampo: 0,67 (0,27-0,92)

Tabla 4: Tabla de resultados del QUADAS-2.

Estudio	Probabilidad de sesgos				Preocupación sobre la aplicabilidad de los resultados		
	Selección de los individuos	Prueba índice	Prueba de referencia	Flujo y tiempos	Selección de los pacientes	Prueba índice	Prueba de referencia
Bruun et al. (2019)	☹	☹	☺	☺	☺	☺	☺
Clerx et al. (2013)	☹	☹	☺	☺	☺	☺	☺
Douaud et al. (2013)	?	☹	?	?	☺	☺	☺
Eckerström et al. (2013)	☹	?	☺	☺	☺	☺	☺
Ferrari et al. (2019)	?	☹	☺	☺	☺	☺	☺
Frölich et al. (2017)	?	☹	☹	☺	☺	☺	?
Gaser et al. (2013)	☹	?	☺	☺	☺	☺	☺
Khan et al. (2014)	☹	?	☺	☺	☺	☺	☺
Ledig et al. (2018)	☹	☹	☺	☺	☺	☺	☺
Monge et al. (2013)	☹	☺	☺	☺	☺	☺	☺
Moscoso et al. (2019)	☹	☹	☺	?	☺	☺	☺
Nesteruk et al. (2016)	☹	☹	?	☺	☺	☺	☺
Ong et al. (2014)	☹	?	☺	☺	☺	☺	☺
Orso et al. (2019)	☺	☹	?	☺	☺	☺	☺
Pereira et al. (2014)	☹	☺	☺	☺	☺	☺	☺
Platero et al. (2019)	☹	☹	?	?	☺	☺	☺
Popescu et al. (2020)	☹	☹	☺	☺	☺	☺	☺
Prestia et al. (2013)	☹	☺	?	?	☺	☺	☺
Prieto et al. (2016)	☹	☹	?	☺	☺	☺	☺
Rhodium et al. (2016)	☹	?	?	☺	☺	☺	☺
Rizzo et al. (2019)	☹	☹	☺	☺	☺	☺	☺
Vos et al. (2013)	☹	☹	?	☺	☺	☺	☺
Wood et al. (2016)	☹	?	?	☺	☺	☺	☺

Tabla 5: Tabla de evidencia de GRADE.

Pregunta: ¿Debería usarse la RM cerebral para diagnosticar demencia de forma precoz?

Desenlace	Nº de estudios (Nº de pacientes)	Diseño de estudio	Factores que pueden disminuir la certeza de la evidencia					Effect per 1000 patients tested	QoE de precisión del test	
			Riesgo de sesgo	Evidencia indirecta	Inconsistencia	Imprecisión	Sesgo de publicación	probabilidad pre-test de 7.7%		
verdaderos positivos (pacientes con demencia)	23 estudios 5761 pacientes	estudios de tipo cohorte y caso-control	muy serio ^a	no es serio	serio ^b	muy serio ^c	ninguno	22 a 69	⊕○○○ Muy baja	
Falsos negativos (pacientes incorrectamente clasificados de no tener demencia)								8 a 55		
Verdaderos negativos (pacientes sin demencia)	23 estudios 5761 pacientes	estudios de tipo cohorte y caso-control	muy serio ^a	no es serio	serio ^b	muy serio ^c	ninguno	388 a 868		⊕○○○ Muy baja
Falsos positivos (pacientes clasificados incorrectamente como demencia)								55 a 535		

Prevalencias	7.7%		
--------------	------	--	--

Explicaciones

- a. selección de pacientes y prueba índice con riesgo elevado de sesgo
- b. en algunos estudios existen valores de sensibilidad y especificidad más extremos
- c. los intervalos de confianza 95% son muy amplios

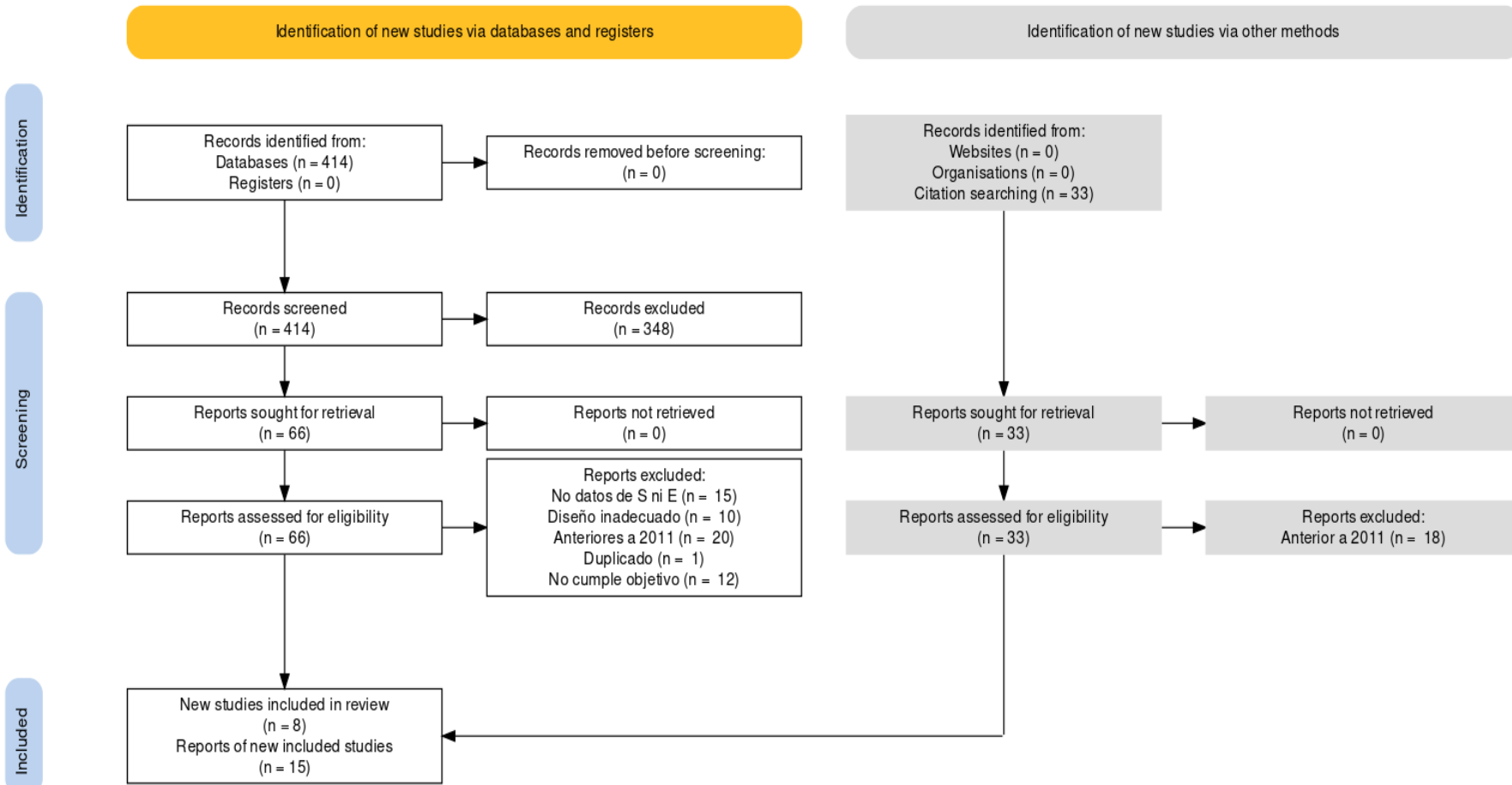


Figura 1: Diagrama de flujo PRISMA sobre la selección de estudios.

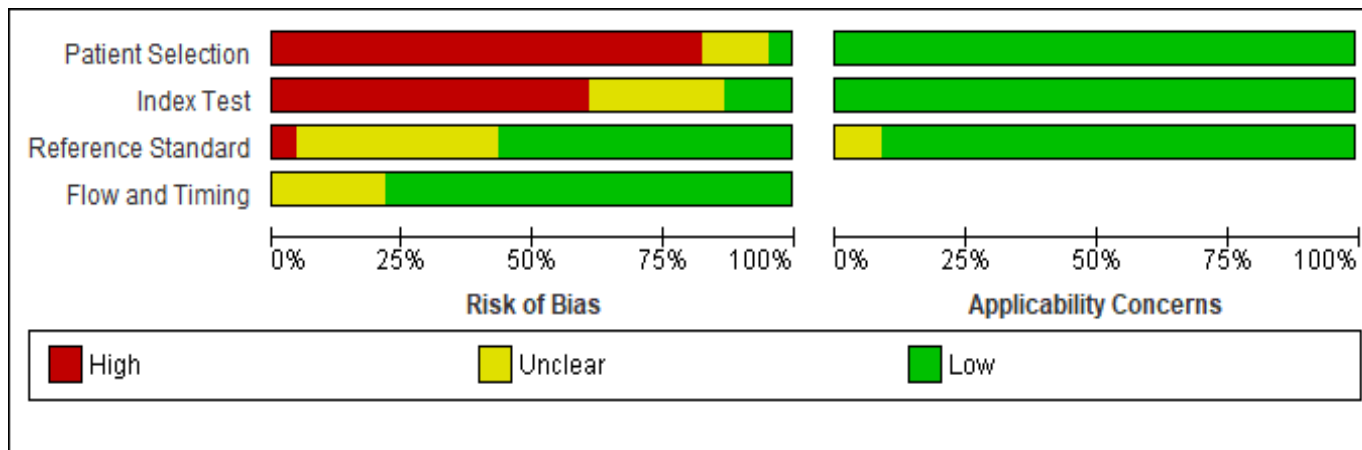


Figura 2. Gráfica QUADAS-2 de los estudios incluidos.

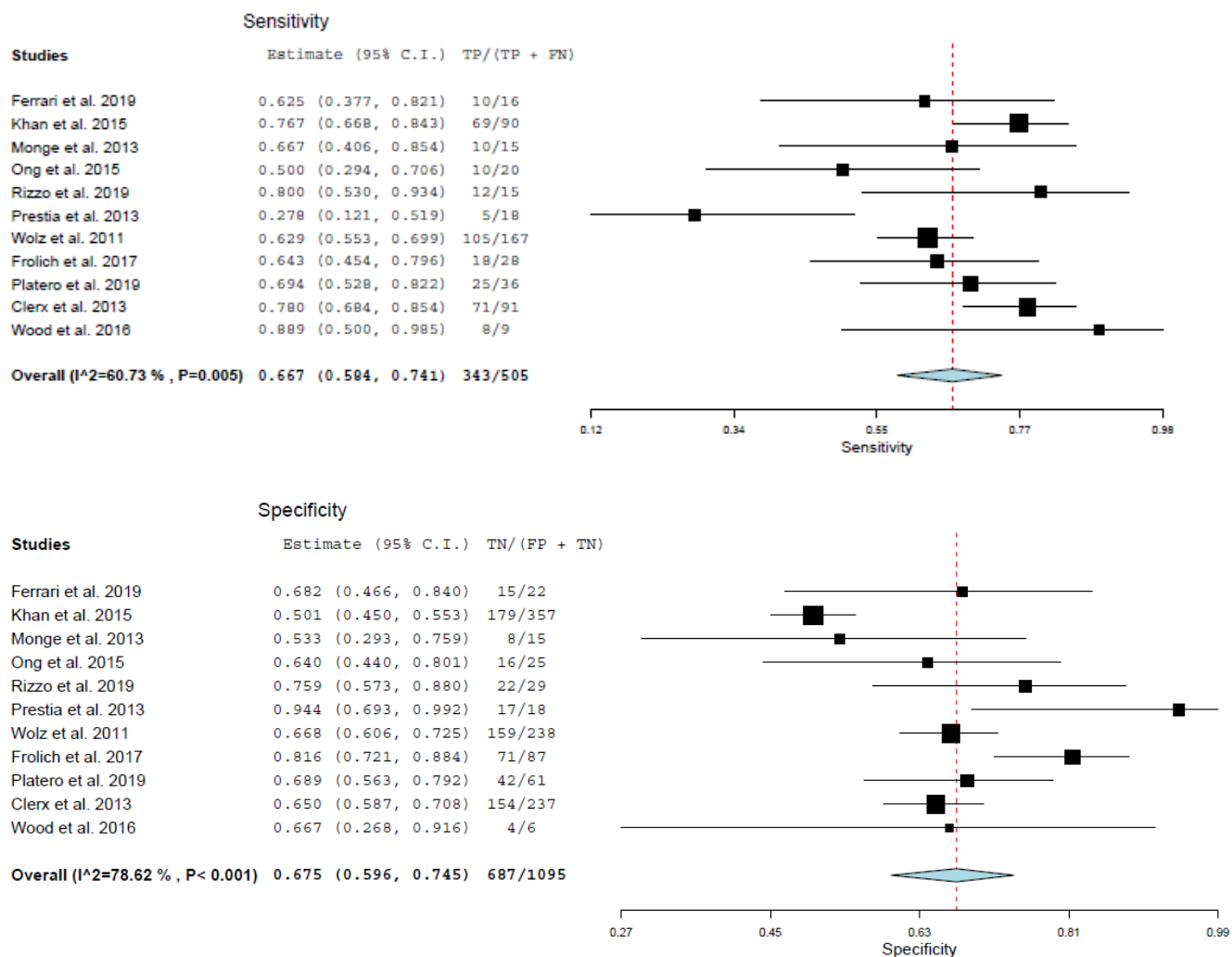


Figura 3. Metanálisis de los estudios incluidos con datos disponibles.

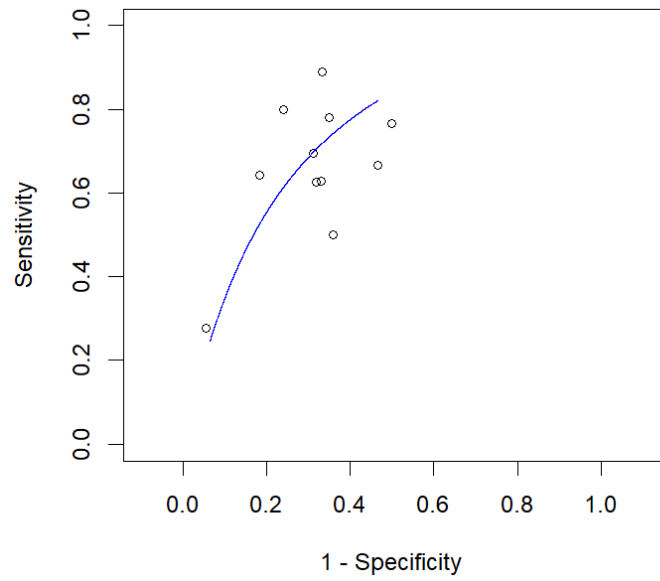


Figura 4. Curva ROC.

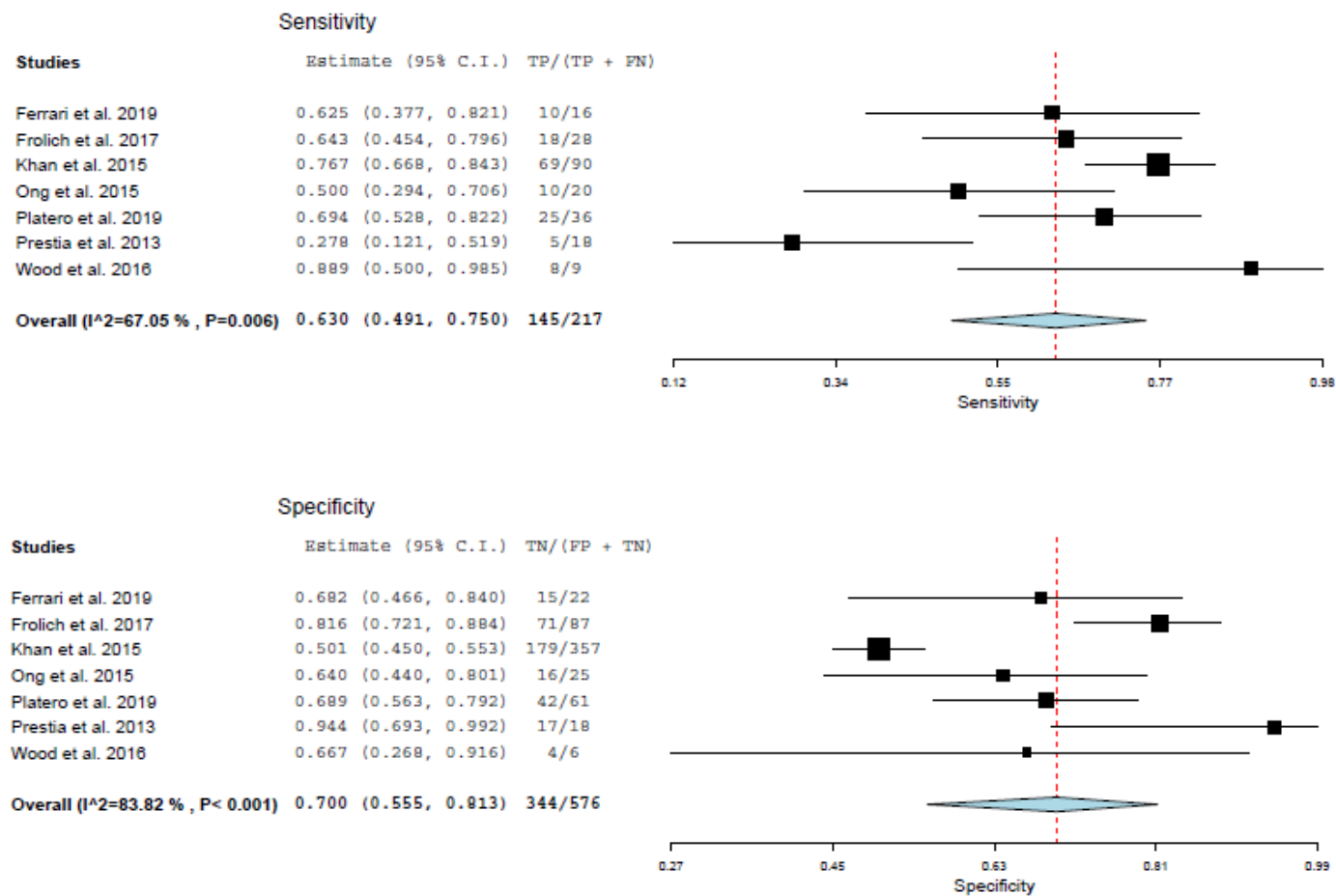


Figura 5. Metanálisis por subgrupo (volumen de hipocampo con medición automatizada).

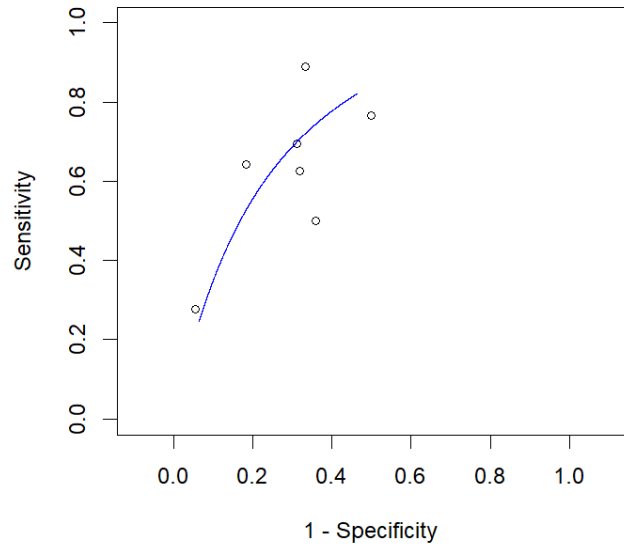


Figura 6. Curva ROC del subgrupo (volumen de hipocampo con medición automatizada).

ACTIVIDAD DESARROLLADA E IMPLICACIÓN DE LA ESTUDIANTE

Como estudiante del máster de Salud Pública, he realizado esta tesina de fin de máster dentro del proyecto de la Agencia de Evaluación de Tecnologías Sanitarias para la realización de una guía de práctica clínica sobre demencias, con el objetivo de actualizar la última existente del año 2011. Se mantuvo una primera reunión con el equipo de tutoras para discutir posibles hipótesis y metodologías del trabajo. Al entrar en este proyecto, ya se encontraba realizado el protocolo del estudio y la estrategia de búsqueda, con lo que mi contribución empezó a partir de la selección de los estudios que cumplían los criterios de inclusión, en un primer momento por título y resumen y posteriormente por texto completo. Más tarde, empecé a analizar la calidad de los estudios incluidos con la herramienta QUADAS-2 y a extraer la información más relevante de cada artículo. Cuando disponía de toda la información, realicé e interpreté el metanálisis (el global y posteriormente el análisis por subgrupo). Igualmente, realicé la tabla de evidencia de GRADE después de mantener una reunión con el equipo de tutoras donde me enseñaron a utilizar dicha herramienta. Finalmente, el manuscrito ha sido realizado por mí con la supervisión y revisiones repetidas por parte del equipo de tutoras, que han ido sugiriendo modificaciones y aclaraciones para obtener el resultado final más óptimo.

ANEXOS

Anexo I: Lista de comprobación PRISMA 2020.

Section and Topic	Item #	Checklist item	Location where item is reported
TITLE			
Title	1	Identify the report as a systematic review.	0
ABSTRACT			
Abstract	2	See the PRISMA 2020 for Abstracts checklist.	
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of existing knowledge.	6-7
Objectives	4	Provide an explicit statement of the objective(s) or question(s) the review addresses.	7
METHODS			
Eligibility criteria	5	Specify the inclusion and exclusion criteria for the review and how studies were grouped for the syntheses.	8
Information sources	6	Specify all databases, registers, websites, organisations, reference lists and other sources searched or consulted to identify studies. Specify the date when each source was last searched or consulted.	
Search strategy	7	Present the full search strategies for all databases, registers and websites, including any filters and limits used.	
Selection process	8	Specify the methods used to decide whether a study met the inclusion criteria of the review, including how many reviewers screened each record and each report retrieved, whether they worked independently, and if applicable, details of automation tools used in the process.	8
Data collection process	9	Specify the methods used to collect data from reports, including how many reviewers collected data from each report, whether they worked independently, any processes for obtaining or confirming data from study investigators, and if applicable, details of automation tools used in the process.	8
Data items	10a	List and define all outcomes for which data were sought. Specify whether all results that were compatible with each outcome domain in each study were sought (e.g. for all measures, time points, analyses), and if not, the methods used to decide which results to collect.	8-9
	10b	List and define all other variables for which data were sought (e.g. participant and intervention characteristics, funding sources). Describe any assumptions made about any missing or unclear information.	
Study risk of bias assessment	11	Specify the methods used to assess risk of bias in the included studies, including details of the tool(s) used, how many reviewers assessed each study and whether they worked independently, and if applicable, details of automation tools used in the process.	9
Effect measures	12	Specify for each outcome the effect measure(s) (e.g. risk ratio, mean difference) used in the synthesis or presentation of results.	8-9

Section and Topic	Item #	Checklist item	Location where item is reported
Synthesis methods	13a	Describe the processes used to decide which studies were eligible for each synthesis (e.g. tabulating the study intervention characteristics and comparing against the planned groups for each synthesis (item #5)).	8
	13b	Describe any methods required to prepare the data for presentation or synthesis, such as handling of missing summary statistics, or data conversions.	
	13c	Describe any methods used to tabulate or visually display results of individual studies and syntheses.	8-9
	13d	Describe any methods used to synthesize results and provide a rationale for the choice(s). If meta-analysis was performed, describe the model(s), method(s) to identify the presence and extent of statistical heterogeneity, and software package(s) used.	9
	13e	Describe any methods used to explore possible causes of heterogeneity among study results (e.g. subgroup analysis, meta-regression).	9
	13f	Describe any sensitivity analyses conducted to assess robustness of the synthesized results.	
Reporting bias assessment	14	Describe any methods used to assess risk of bias due to missing results in a synthesis (arising from reporting biases).	
Certainty assessment	15	Describe any methods used to assess certainty (or confidence) in the body of evidence for an outcome.	9
RESULTS			
Study selection	16a	Describe the results of the search and selection process, from the number of records identified in the search to the number of studies included in the review, ideally using a flow diagram.	10
	16b	Cite studies that might appear to meet the inclusion criteria, but which were excluded, and explain why they were excluded.	10
Study characteristics	17	Cite each included study and present its characteristics.	13-17
Risk of bias in studies	18	Present assessments of risk of bias for each included study.	37
Results of individual studies	19	For all outcomes, present, for each study: (a) summary statistics for each group (where appropriate) and (b) an effect estimate and its precision (e.g. confidence/credible interval), ideally using structured tables or plots.	13-17
Results of syntheses	20a	For each synthesis, briefly summarise the characteristics and risk of bias among contributing studies.	29-39
	20b	Present results of all statistical syntheses conducted. If meta-analysis was done, present for each the summary estimate and its precision (e.g. confidence/credible interval) and measures of statistical heterogeneity. If comparing groups, describe the direction of the effect.	42, 44
	20c	Present results of all investigations of possible causes of heterogeneity among study results.	13-17
	20d	Present results of all sensitivity analyses conducted to assess the robustness of the synthesized results.	
Reporting biases	21	Present assessments of risk of bias due to missing results (arising from reporting biases) for each synthesis assessed.	

Section and Topic	Item #	Checklist item	Location where item is reported
Certainty of evidence	22	Present assessments of certainty (or confidence) in the body of evidence for each outcome assessed.	39
DISCUSSION			
Discussion	23a	Provide a general interpretation of the results in the context of other evidence.	18-20
	23b	Discuss any limitations of the evidence included in the review.	20
	23c	Discuss any limitations of the review processes used.	20
	23d	Discuss implications of the results for practice, policy, and future research.	22
OTHER INFORMATION			
Registration and protocol	24a	Provide registration information for the review, including register name and registration number, or state that the review was not registered.	
	24b	Indicate where the review protocol can be accessed, or state that a protocol was not prepared.	
	24c	Describe and explain any amendments to information provided at registration or in the protocol.	
Support	25	Describe sources of financial or non-financial support for the review, and the role of the funders or sponsors in the review.	
Competing interests	26	Declare any competing interests of review authors.	
Availability of data, code and other materials	27	Report which of the following are publicly available and where they can be found: template data collection forms; data extracted from included studies; data used for all analyses; analytic code; any other materials used in the review.	

Este estudio no ha precisado la autorización de un comité de ética de investigación, dado que se trata de una revisión sistemática en la que no se ha precisado la colaboración de pacientes ni su autorización para realizar apartado alguno del presente trabajo.