Ocular biomarkers for Alzheimer disease dementia: an umbrella review of systematic reviews and meta-analyses

Costanzo, E., Lengyel, I., Parravano, M., Biagini, I., Veldsman, M., Badhwar, A., Betts, M., Cherubini, A., Llewellyn, D. J., Lourida, I., MacGillivray, T., Rittman, T., Tamburin, S., Tai, X. Y., & Virgili, G. (2023). Ocular biomarkers for Alzheimer disease dementia: an umbrella review of systematic reviews and meta-analyses. *JAMA Ophthalmology*, *141*(1), 84-91. https://doi.org/10.1001/jamaophthalmol.2022.4845

Published in:
JAMA Ophthalmology

Document Version:
Peer reviewed version

Queen's University Belfast - Research Portal:
Link to publication record in Queen's University Belfast Research Portal

Ocular biomarkers for Alzheimer's Disease Dementia: An Umbrella Review of Systematic Reviews and Meta-analyses

Review team: Eliana Costanzo[1], Imre Lengyel[2], Mariacristina Parravano[1], Ilaria Biagini[3], Michele Veldsman[4], AmanPreet Badhwar[5,6], Matthew Betts[7,8,,9], Antonio Cherubini[10], David Llewellyn[11], Ilianna Lourida[11], Tom MacGillivray[12], Timothy Rittman[13], Stefano Tamburin[14], Xin You Tai[15], Gianni Virgili*[3,16]

1. IRCCS - Fondazione Bietti, Rome, Italy
2. Wellcome-Wolfson Institute for Experimental Medicine, Queen's University Belfast (UK)
3. Department NEUROFARBA, University of Florence, Italy
4. Department of Experimental Psychology, University of Oxford, Oxford, UK
5. Department of Pharmacology and Physiology, University of Montreal, Montreal, Canada
6. Centre de recherche de l'Institut Universitaire de Geriatrie, Montreal, Canada
7. Institute of Cognitive Neurology and Dementia Research, Otto-von-Guericke University Magdeburg, Magdeburg, Germany
8. German Center for Neurodegenerative Diseases (DZNE), Magdeburg, Germany
9. Center for Behavioral Brain Sciences, University of Magdeburg, Magdeburg, Germany
10. Geriatria, Accettazione geriatrica e Centro di ricerca per l'invecchiamento. IRCCS INRCA Ancona (Italy)
11. College of Medicine and Health, University of Exeter, Exeter, UK
12. Centre for Clinical Brain Sciences, University of Edinburgh
13. Department of Clinical Neurosciences, University of Cambridge
14. Department of Neurosciences, Biomedicine and Movement Sciences, University of Verona, Italy
15. Nuffield Department of Clinical Neurosciences, University of Oxford
16. Centre for Public Health, Queens University Belfast (UK)

*Corresponding author: Gianni Virgili,
Centre for Public Health, Queen’s University Belfast, Belfast (UK)

Institute of Clinical Science, Block A, Royal Victoria Hospital, Belfast BT12 6BA

Phone: +44(0) 28 9097 6350

Email: G.Virgili@qub.ac.uk

Words' count: 3290 words

Key point

- **Question**: What is the diagnostic accuracy of ocular biomarkers for early diagnosis of Alzheimer's disease (AD), as investigated in systematic reviews?
- **Finding**: Fourteen SRs and meta-analyses were included. OCT peripapillary RNFL thickness, OCTA FAZ measurement and prosaccade latency of saccadic eye movements were extensively investigated and yielded only moderate accuracy. Antisaccade error showed the best accuracy in a smaller number of trials.
- **Meaning**: Ocular biomarkers showed poor to moderate diagnostic accuracy for detection of AD in cross-sectional studies. Longitudinal studies are needed on whether changes in these parameters could yield better predictions of AD onset.

Abstract

Importance: Several ocular biomarkers have been proposed for the early detection of Alzheimer's disease (AD) and Mild Cognitive Impairment (MCI), particularly fundus photography, Optical Coherence Tomography (OCT) and OCT-angiography (OCTA).

Objective: Umbrella review of systematic reviews (SRs) on this topic.

Data sources: MEDLINE, EMBASE and PsycINFO from January 2000 to November 2021.

Study selection: Title and abstracts were screened independently by two reviewers. No language restriction was used.

Data extraction and synthesis: The Risk Of Bias in Systematic reviews (ROBIS) tool was used to assess quality. Diagnostic accuracy was presented descriptively and graphically.

Main outcomes and measures: The Area under the Curve (AUC) was derived from Standardised Mean Difference (SMD).

Results: From the 591 titles, we included 14 SRs (median 14 studies in each review, range 5 to 126). OCT and/or OCTA were considered in 9 reviews, fundus photography in 3 SRs and saccadic eye movements (SEM) in 2 SRs, and one review included all tests. Only four reviews were at 'low' risk of bias on all ROBIS domains. Most provided no or limited discussion of the target population and test role in the clinical pathway. Most primary studies were of a case-control nature, but did not account for the risk of overestimating diagnostic performance with this design.

The imaging-derived parameters with most evidence for detecting AD compared to healthy controls (HC), were: OCT peripapillary retinal nerve fiber layer thickness (38 studies, 1883 AD patients and 2510 HC; AUC=0.70; $p<0.001$); OCTA foveal avascular zone (5 studies, 177 AD patients and 371 HC; AUC=0.73; 0.500, 0.893); SEM prosaccade latency (30 studies, 651 AD/MCI patients and 771 HC; AUC=0.64; 0.584, 0.687). Antisaccade error was investigated in fewer studies (12, 424 AD/MCI patients and 382 HC), and yielded the best accuracy (AUC=0.79; 0.695, 0.880).

Conclusions and Relevance: Our umbrella review has highlighted limitations in design and reporting of the existing research on ocular biomarkers for diagnosing AD. Parameters with the best evidence showed poor to moderate diagnostic accuracy in cross-sectional studies. We suggest that future longitudinal studies should investigate whether changes in OCT and OCTA measurements over time can yield accurate predictions of AD onset.

Introduction

The early and accurate diagnosis of Alzheimer's disease (AD) and related-dementias is critical to clinical disease management and for stratification in trials of disease modifying therapies.[1,2]

Biomarkers that directly measure neuropathology are increasingly used to support diagnosis.[3] However, many of such neuroimaging and cerebrospinal fluid biomarkers are invasive, expensive or not widely available.[4] Ocular biomarkers hold the potential to provide objective, affordable and widely available measurements that reflect the underlying neuropathology of dementia, [4,5] also non-invasively using devices as optical coherence tomography (OCT), OCT-angiography (OCTA) and retinal colour fundus photograph,[6] and also with the analysis of saccadic eye movements.

Here we review the progress in translating such ocular biomarkers into clinical practice for the diagnosis of dementia, in particular AD and Mild Cognitive Impairment (MCI), which may represent its earliest stage.

The retina is part of the central nervous system with direct connection with different brain areas; it has even been suggested that axons of the optic nerve facilitate the transport into the brain of amyloid precursor protein (APP) created in retinal ganglion cells (RGCs).[4] Many brain areas are involved in oculomotor control and neurodegenerative diseases present with oculomotor and saccadic abnormalities.[7] Patients with AD and other neurodegenerative diseases may report decreased vision, visual field changes, visual hallucinations and other visual symptoms.[8] The ocular changes that occur in AD include abnormal pupillary reaction, decreased contrast sensitivity, loss of RGCs and retinal nerve fibre layer (RNFL), peripapillary atrophy, and retinal thinning, tortuosity of blood vessels and deposition of beta-amyloid in the retina.[9]

Given these neurobiological associations between the eye and the brain, efforts are underway to establish ocular biomarkers of dementia. A number of small reviews have attempted to synthesise

specific aspects of the literature. However, the field is still developing and there are potential risks of bias and small sample size that could hamper efforts, so it remains unclear which methods are most promising. The most widely studied parameters from ophthalmic techniques are retinal nerve fiber layer (RNFL) thickness, ganglion cells-inner plexiform layer (GC-IPL) complex, foveal avascular zone (FAZ) area, vessel density (VD) and perfusion density (PD), that are altered in dementia.[4,6,10-21]

In addition, the retinal fundus imaging might be used as a surrogate for brain vascular changes,[4,6,22-24] revealing subtle modifications in the cerebral vasculature related to preclinical stages of neurodegenerative diseases, such as arteriolar narrowing and venular widening or e.g., the low complexity or density of retinal vessels expressed as reduction of the fractal dimensions(FD).[23,24] These changes might reflect subtle modifications in the cerebral vasculature related to preclinical stages of neurodegenerative diseases, suggesting that retinal fundus imaging might be used as a surrogate for brain vascular changes.[4,6,22,24,25]

Meta-epidemiologic studies have shown that multiple systematic reviews on the same topic have variable quality and their results and conclusions may be only partly overlapping, for to the time span of literature searches, different inclusion criteria and analytic approaches, the use of subgroup analyses and for the interpretation of results. [26-29]

The aim of this umbrella review was to summarise the findings of published systematic reviews on the diagnostic performance of ocular biomarkers for detecting dementia and the prognostic significance for assessing the risk of conversion of mild cognitive impairment (MCI) to dementia.

Methods

This umbrella systematic review, or overview of reviews, is reported following PRISMA guideline.[30] (see PRISMA checklist in the Supplement). A protocol was developed following discussion with topic and methods experts and is registered with PROSPERO (PROSPERO 2021: CRD42021287196). Medline, EMBASE and PsycINFO were searched from January 2000 to November 2021 and the references of included reviews were also searched. Full details of the methods for conducting the review and search strategy are reported in the Supplement.

Systematic reviews were included if they investigated the diagnostic accuracy of ocular biomarkers to detect AD, all-cause dementia or MCI against established clinical criteria or clinical judgement. Eligible reviews assessed the accuracy of any technique related to eyes or vision, including imaging, functional assessment, pathology, or laboratory testing in secondary care or specialist diagnostic services settings such as memory clinics. We excluded studies using questionnaires, even if vision-specific domains were adopted, since we aimed to assess objective biomarkers addressing a single physiologic trait. The methodological quality of the included reviews was assessed using the Risk of Bias in Systematic Reviews (ROBIS) tool.[31]

Primary measures of effect were sensitivity, specificity and area under the curve (AUC). Since none of these were available, we extracted continuous data for each biomarker, with a preference for Standardised Mean Differences (SMD) between diseased and non-diseased groups to account for differences in the measurement scale or calibration of different devices/techniques, using a conversion tables to infer the AUC as a measure of diagnostic accuracy.[32]. If the SMD was not available, we presented Mean Differences (MDs). Based on previous literature,[33] we considered AUC=0.70 as the threshold between poor and acceptable diagnostic accuracy, and 0.80 as the threshold between acceptable and excellent accuracy. However, we suggest an AUC 0.80 should be considered useful for prognostic factors, but still moderate for diagnostic accuracy. In fact, the associated sensitivity and specificity pair with AUC=0.80 at the inverse diagonal (the Q-point) are

104 only 0.72 for a symmetric Receiver Operator Curve (ROC); even a nominally excellent AUC=0.90
105 would generate a moderately good sensitivity and specificity pair at 0.82, meaning that 1 in 5 test
106 positives are false positives and 1 in 5 test negatives are false negatives.

107 We found a large number of parameters for the two main methods, OCT and OCTA. We
108 therefore selected the three parameters with the largest number of studies to provide the best
109 estimation of performance. When the third ranking was shared for more parameters, we presented
110 all of them. When a parameter was collected in different retinal locations, e.g., parafovea, perifovea,
111 fovea, or whole, we followed this order of selection.[34]

112

113 **Results**

114 After duplicates were manually removed, the search retrieved 591 references, of which 32
115 systematic reviews were obtained in full text. Eighteen of these were excluded for one or more of
116 the following reasons: published before 2016 (7 reviews)[35-41]; used surrogates of dementia as a
117 target condition (1 review)[22]; used ocular conditions as biomarker, rather than test technique results
118 (2 reviews)[42,43]; absence of meta-analyses (10 reviews).[4,5,9,11,14,21,22,44-46] Five reviews published
119 before 2016 reported on OCT[35-38,41] and 2 on retinal microvascular changes.[39,40] Consistent with our
120 expectations, the number of included studies in these reviews was much lower than that in the most
121 recent SRs on the same techniques.

122 Fourteen included articles were systematic reviews with meta-analyses. Additional studies
123 reported quantitative individual study results with no meta-analysis, and narrative study results
124 (Fig.1). Due to limitations in reporting, data were not extracted and presented for the latter two
125 groups of reviews.

126 Analysed reviews were published between 2016 and 2021, and included 5 to 126 studies
127 (median 14 studies). Included techniques were OCT (n=7), OCTA (n=5), fundus imaging (n=3),

and saccadic eye movements (n=2). One review also reported Scanning Laser Ophthalmoscopy (SLO)/polarimetry results, which were not considered since this device is no longer commercially available. Three reviews focused on AD as the target condition,[18,19,23], four on MCI,[13,15,16,24] and seven included both[6,10,12,17,20,47,48].

Regarding measures of effect, eight studies used SMD to measure effect sizes,[6,10,12,15,16,19,47,48] and six used the MD,[13,17,18,20,23,24] with one review also using the ratio of means to pool data for meta-analyses.[6]

Regarding the risk of bias in studies, three reviews used Quality Assessment of Diagnostic Accuracy Studies (QUADAS) or QUADAS 2,[10,12,16] four used the Newcastle Ottawa Scale (NOS),[17-19,24] one used the Joanna Briggs Institute critical appraisal tools,[6] one used Risk Of Bias In Non-randomized Studies Of Interventions (ROBINS-I),[47] and the other did not adopt validated instruments.[13,20,23,48] One review used Risk-of-bias VISualization (ROBVIS),[15] a visualisation tool, and used questions designed for randomised controlled trials, though some additional items were added.

eTable 1 in the Supplement shows the main characteristics of 14 included systematic reviews which reported quantitative estimates, and the risk of bias assessment with ROBIS is detailed in eTable 2 in the Supplement.

Overview of all parameters

Several parameters were reported in the included reviews. As explained in the Methods, for each technique we selected the three parameters with the largest number of studies. These were for OCT; pRNFL[6,10,12,13,15,16], macular RNFL (mRNFL)[6] and GC-IPL complex[6,10,15,16] thickness; for OCTA: FAZ area[6,17-20], PD/VD in the superficial capillary plexus (SCP)[6,17-20] and in the deep capillary plexus (DCP)[6,17,18]; for fundus camera: central retinal venular equivalent caliber (CRVE)[6,23,24], FD total, arterial and venous[6,23,24]; for SEM: prosaccade latency, antisaccade latency

and antisaccade error[47,48]. For each parameter we analysed the comparisons between AD vs HC, AD vs MCI and MCI vs HC.

All the results for each study and for the different techniques, parameters and comparisons are extensively detailed by reporting all the SMD/MD values, p-values/95%CI and corresponding AUC, in eTable 3 in the Supplement. Of note, perfusion density (PD) was expressed as %, but vessel density (VD), which has a similar clinical significance was used and expressed as mm^2 in four reviews on SCP[6,17-20]. No review used adjusted estimates of test accuracy, e.g., accounting for age-related changes, but they reported studies had matched cases and controls by age and gender.

In the following paragraphs we report on the three parameters with the best performance for each technique. These are also presented as forest plots (OCT in Fig.2, OCTA in Fig.3, and SEM in Fig.4).

OCT pRNFL

Mean pRNFL thickness was reported in the largest number of studies in three systematic reviews (Fig.2; eTable 3 in the Supplement). The pooled SMD for detecting AD vs HC was -0.723, corresponding to an AUC=0.70 in the most extensive review (38 studies),[6] which was at low risk of bias. Two reviews[6,12] pooled data comparing AD with MCI for this parameter. The review with low risk of bias and more studies[6] found the SMD was -0.140 (p=0.064), corresponding to an AUC=0.54.

Six systematic reviews[6,10,12,13,15,16] pooled data comparing MCI with HC for this parameter. The review with low risk of bias and more studies[6] reported a pooled SMD of -0.324 (p<0.001), corresponding to an AUC=0.59, with high heterogeneity.

OCTA FAZ area

Five reviews [6,17-20] reported on the FAZ in AD vs HC, with 3 to 7 studies each, and found concordant estimates of difference (Fig. 3; eTable 3 in the Supplement). The two reviews using the

SMD found SMD=0.87 (95%CI -0.01, 1.76), with imprecision crossing no difference and with high heterogeneity.[6,19] The AUC corresponding to this point estimate was 0.73.

Two reviews,[6,20] including two studies each (one overlapping study) compared AD and MCI. One review[6] found an SMD of 0.52 (95%CI -0.38, 1.41) with considerable imprecision and heterogeneity. The corresponding AUC point estimate was 0.64.

Three reviews, [6,17,20] including 3 or 4 studies each, included MCI patients and HC. The review using the SMD found a pooled value of 0.83 (95%CI -0.19, 1.85), with considerable imprecision and heterogeneity, corresponding to an AUC=0.72.

Saccadic eye movements

Saccadic eye movements were explored in two systematic reviews with meta-analysis [47,48] which reported significant alterations in AD and MCI patients (Fig.4; eTable 3 in the Supplement). The parameters with the highest number of studies were: prosaccade latency, antisaccade latency and antisaccade error (investigated with gap, overlap or unspecified conditions). Both reviews combined studies on AD with a minority of studies on MCI, all compared with HC. One review[47] compared prosaccade latency in gap and overlap conditions finding SMD=0.30 and 0.50, respectively (95%CI -0.05, -0.01). The antisaccade latency and antisaccade error were reported in gap condition by Opwonya et al.,[47] finding SMD=0.44 and 1.16, with AUC=0.62 and 0.79, respectively and were reported in undefined condition by Kahana Levy et al.,[48] finding SMD=0.37 and 1.30, with AUC=0.60 and 0.82, respectively.

The parameter with most studies (25 to 30) was prosaccade latency (investigated with gap, overlap or unspecified condition) for which AUCs were 0.58 to 0.64 with sufficient precision. Antisaccade latency yielded similar results to prosaccade latency.

198 Among those selected, the parameter with the best performance was antisaccade error (with
199 gap or unspecified condition) in 8-12 studies, with AUC=0.79, which, however, was imprecisely
200 estimated.

201 **Other biomarkers**

202 Ge et al.[6] reported narratively on studies using retinal blood flow and electrophysiological
203 biomarkers (multifocal or pattern ERG, not presented here). They also reported narratively on
204 several other candidate biomarkers assessed in individual studies, (e.g., retinal blood flow with
205 different instruments, retinal arterial and venous oxygen saturation).

206 Amyloid deposition in one hyperspectral imaging study was found to have an AUC=0.82
207 (0.67-0.97) for 15 AD or MCI subjects compared to HC.[49] A second study confirmed this finding,
208 but estimates were not available.[50] Other imaging studies used Blue-Peak autofluorescence.[25] Ge
209 2021[6] also reported anecdotally on studies using reflective features related to the Müller cells and
210 macular pigment optical density. All these additional biomarkers have not been included in our
211 analyses because no meta-analyses were performed in systematic reviews.

212

213 **Discussion**

214 Our umbrella review found several systematic reviews that have assessed the diagnostic
215 performance of ocular biomarkers, especially OCT and OCTA to detect AD and MCI.

216 We found that the most widely investigated eye biomarkers were pRNFL thickness
217 measured with OCT, FAZ area evaluated with OCTA, and prosaccade or antisaccade latency or
218 error for saccadic eye movements.

219 In two reviews at low risk of bias,[6,10] AUCs were low (about 0.70) for the most widely
220 studied imaging parameters (pRNFL and FAZ area). Results were variable and generally modest
221 and imprecisely estimated for all other OCT parameters. In order to highlight the limited usefulness

of such AUC values, we calculated that sensitivity and specificity pairs would be equal at 64.4% at the inverse diagonal (Q point), and that the specificity at 95% would lead to only 19% sensitivity, suggesting a modest performance. This is consistent with the finding by Ge et al.,[6] who reported the ratio of means of pRNFL between AD and HC was 0.93, i.e., only a 7% difference in average RNFL thickness between the two groups.

Saccadic eye movements were investigated in a large number of studies included in two reviews, one of which was low risk of bias.[47] Prosaccade and antisaccade latency yielded AUC values below 0.70. The parameter with a better performance was antisaccade error (AUC=0.79), suggesting that it should be considered for further research.

Other parameters were CRVE and FD evaluated with retinal photography, yielding imprecise estimates of poor diagnostic performance. Other vascular parameters as Central Retinal Arteriolar equivalent caliber (CRAE), arterial-to-vein ratio (AVR), branching coefficient and tortuosity have been reviewed in AD/MCI population,[6] yielding lower diagnostic accuracy, and have not been included in the present analysis.

Several other reviews could have been included but did not provide data. Alber et al.[4] and Mahajan et al.[9] were broad-spectrum but primarily narrative reviews and quantitative data were not considered for meta-analysis. A number of reviews[5,11,14,44] were systematic reviews of OCT data in several neurologic conditions with no meta-analysis. Wu et al.[24] provided quantitative data with OCT and OCTA and reported differences between AD and MCI but did not conduct a meta-analysis.

Other reviews included a broad range of diseases, Zhang et al.[21] investigated OCTA in patients with small vessel disease, including stroke and dementia. Lemmens et al.[22] found a significant association between decreased FD and neurodegenerative disease. Youssef et al.[45] reported variation in electroretinography findings in several psychiatric disorders.

Overall, several systematic reviews have been produced just in the last five years, the largest of which including over 100 studies. None of these reviews could generate inference that can be used in clinical practice, and only Ge 2021[6] declared they would synthesis these effect measure in the Method section.

No included review framed the research question in a clinical pathway, which is key in diagnostic test accuracy reviews, since patient's prior testing and disease spectrum, often related to study setting and test user profile, are known to influence accuracy and also its translation into practice.[51]

The quality of the included studies was not taken into account to generate conclusions in the majority of the reviews, neither was the heterogeneity of study findings. Overall, only four[6,10,16,47] out of 14 reviews were at low risk of bias for all ROBIS domains.

A further level of complexity is that studies included different devices, using different scales for the measurement of the same biomarker. The comparison between different devices is difficult,[52] and many reviews used SMD as a measure of effect to overcome this problem. The SMD metrics for therapeutic clinical significance commonly assumes small effects if SMD<0.5 and large effects if SMD>0.8. Much larger differences are meaningful in diagnostic accuracy, where an SMD=1.81 equals an AUC=0.90, i.e., a moderately good sensitivity=specificity=0.82 at Q-point.

The modest diagnostic performance we found for the two main techniques of greater interest, OCT and OCTA, does not mean these have no use in dementia investigation. First of all, OCTA has not yet been extensively used. Glaucoma research with OCT has shown that, while cross-sectional accuracy is poor due to interindividual variation in normal subjects, the change in OCT parameters over time in the same individual can be helpful for detection.[53] Therefore, longitudinal studies on AD and MCI development might result in better diagnostic accuracy compared to cross-sectional detection. Finally, even biomarkers with weak diagnostic performance could be used in artificial intelligence-based algorithms for case finding or prediction, provided that they contribute independently from other variables.

Open issues include testing the sensitivity and specificity of ocular biomarkers compared to more robust and widely accepted biomarkers (e.g., CSF, PET) of amyloid, tau, and neuronal loss. This piece of information, compared to diagnostic yield in populations selected on the basis of a clinical diagnosis would help clarifying whether the weak diagnostic performance is real or due to loose clinical diagnostic criteria across studies. Saccadic eye movements also merit further investigation, possibly regarding antisaccade error.

We suggest that future studies on eye-imaging biomarkers in dementia should clearly report: diagnostic criteria for dementia/MCI, whether the diagnosis is clinical or supported by biomarkers, the staging of the disease, longitudinal follow-up (when available), post-mortem neuropathological confirmation. Most importantly, an a-priori diagnostic question should be framed instead of an exploratory design of the study and follow-up data should be obtained in community-based samples, recruiting individuals in the lowest range of normative data for OCT and/or OCTA parameters. In fact, Dong et al.[54] found that this group provides the best opportunity to directly visualize subclinical changes in neurosensory and vascular tissue that are very likely to occur in many prevalent chronic diseases and new methods for detecting these changes are in development.[55]

A strength of our overview of reviews was the broad and inclusive approach, encompassing all techniques and conditions related to dementia that had been reported in systematic reviews, based on a research protocol and using a well-established risk of bias tool for systematic reviews (ROBIS), which accommodates different study designs. A limitation of our review is that we did not formally assess study overlap between systematic reviews, and we did not attempt a meta-analysis of original studies, nor did we assess their individual quality. Instead, we presented the quality of the included systematic reviews, meaning that we only assessed whether optimal evidence synthesis methods were used. Nonetheless, we think the consistently moderate diagnostic performance found for ocular biomarkers in all reviews means that results would not change if such a large effort was undertaken. Despite this, the evaluation of non-invasive ocular biomarkers for

296 early AD diagnosis could be useful in artificial intelligence models or in large-scale screening
297 programs selecting patients for specialist referrals.

298

299 **Conclusions**

300 Our overview of reviews showed that research on ocular biomarkers for diagnosing AD,
301 summarised in systematic reviews, was abundant in recent years, but poorly reported and of limited
302 clinical significance. Further studies should adopt a formal diagnostic accuracy framework and
303 follow STARD guidance[56] for reporting results. Longitudinal studies should investigate whether
304 changes in OCT and OCTA measurements over time can yield accurate predictions of AD onset.

Acknowledgments

We are grateful to Iris Gordon (Information Specialist, Centre for Public Health, Queens University Belfast, UK) for conducting the searches for this review. The research for this paper was financially supported by the Italian Ministry of Health and Fondazione Roma.

Eliana Costanzo and Gianni Virgili had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis

Declarations of interest: Eliana Costanzo, Ilaria Biagini, Imre Lengyel, Michele Veldsman, AmanPreet Badhwar, Matthew Betts, Antonio Cherubini, David Llewellyn, Ilianna Lourida, Tom MacGillivray, Timothy Rittman, Stefano Tamburin, Xin You Tai and Gianni Virgili have nothing to disclose. Mariacristina Parravano reports personal fees from Allergan, Novartis, Bayer, Roche, Zeiss outside the submitted work.

Funding

The research reported in this publication was granted by the Italian Ministry of Health, under the Aging Network of Italian Research Hospitals (IRCCS). The work was also supported by grant from the Medical Research Council of UK (MR/N029941/1) and an unrestricted grant from OPTOS Plc (IL). The funders had no role in study design, data collection, analysis, decision to publish, or manuscript preparation.

Author contributions:

Eliana Costanzo: conceptualization of the work, data analysis and interpretation of the results, manuscript drafting and final approval. Imre Lengyel: conceptualization of the study, data

interpretation, manuscript drafting and critical revision. Mariacristina Parravano: data analysis and interpretation of results. Ilaria Biagini: data analysis and tables' creation. Michele Veldsman, AmanPreet Badhwar, Matthew Betts, Antonio Cherubini, David Llewellyn, Ilianna Lourida, Tom MacGillivray, Timothy Rittman, Stefano Tamburin, Xin You Tai: critically revision of manuscript. Gianni Virgili: conceptualization and design of the work, data analysis and interpretation of the results, manuscript drafting and final approval.

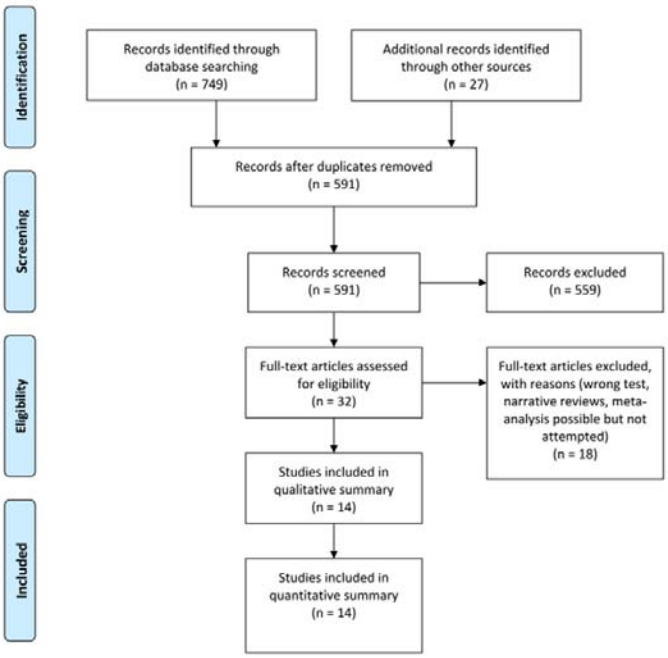
References

1. Livingston G, Huntley J, Sommerlad A, et al. Dementia prevention, intervention, and care: 2020 report of the Lancet Commission. *Lancet*. 08 08 2020;396(10248):413-446. doi:10.1016/S0140-6736(20)30367-6
2. Yiannopoulou KG, Anastasiou AI, Zachariou V, Pelidou SH. Reasons for Failed Trials of Disease-Modifying Treatments for Alzheimer Disease and Their Contribution in Recent Research. *Biomedicines*. Dec 09 2019;7(4)doi:10.3390/biomedicines7040097
3. Soria Lopez JA, González HM, Léger GC. Alzheimer's disease. *Handb Clin Neurol*. 2019;167:231-255. doi:10.1016/B978-0-12-804766-8.00013-3
4. Alber J, Goldfarb D, Thompson LI, et al. Developing retinal biomarkers for the earliest stages of Alzheimer's disease: What we know, what we don't, and how to move forward. *Alzheimers Dement*. 01 2020;16(1):229-243. doi:10.1002/alz.12006
5. Song A, Johnson N, Ayala A, Thompson AC. Optical Coherence Tomography in Patients with Alzheimer's Disease: What Can It Tell Us? *Eye Brain*. 2021;13:1-20. doi:10.2147/EB.S235238
6. Ge YJ, Xu W, Ou YN, et al. Retinal biomarkers in Alzheimer's disease and mild cognitive impairment: A systematic review and meta-analysis. *Ageing Res Rev*. 08 2021;69:101361. doi:10.1016/j.arr.2021.101361
7. Kassavetis P, Kaski D, Anderson T, Hallett M. Eye Movement Disorders in Movement Disorders. *Mov Disord Clin Pract*. Apr 2022;9(3):284-295. doi:10.1002/mdc3.13413
8. Douglas VP, Douglas KAA, Cestari DM. Ophthalmic manifestations of dementing disorders. *Curr Opin Ophthalmol*. Nov 01 2021;32(6):515-520. doi:10.1097/ICU.0000000000000807
9. Mahajan D, Votruba M. Can the retina be used to diagnose and plot the progression of Alzheimer's disease? *Acta Ophthalmol*. Dec 2017;95(8):768-777. doi:10.1111/aos.13472
10. Chan VTT, Sun Z, Tang S, et al. Spectral-Domain OCT Measurements in Alzheimer's Disease: A Systematic Review and Meta-analysis. *Ophthalmology*. 04 2019;126(4):497-510. doi:10.1016/j.ophtha.2018.08.009
11. Chhablani PP, Ambiya V, Nair AG, Bondalapati S, Chhablani J. Retinal Findings on OCT in Systemic Conditions. *Semin Ophthalmol*. 2018;33(4):525-546. doi:10.1080/08820538.2017.1332233
12. den Haan J, Verbraak FD, Visser PJ, Bouwman FH. Retinal thickness in Alzheimer's disease: A systematic review and meta-analysis. *Alzheimers Dement (Amst)*. 2017;6:162-170. doi:10.1016/j.dadm.2016.12.014
13. Knoll B, Simonett J, Volpe NJ, et al. Retinal nerve fiber layer thickness in amnestic mild cognitive impairment: Case-control study and meta-analysis. *Alzheimers Dement (Amst)*. 2016;4:85-93. doi:10.1016/j.dadm.2016.07.004
14. López-de-Eguileta A, Cerveró A, Ruiz de Sabando A, Sánchez-Juan P, Casado A. Ganglion Cell Layer Thinning in Alzheimer's Disease. *Medicina (Kaunas)*. Oct 21 2020;56(10)doi:10.3390/medicina56100553
15. Mejia-Vergara AJ, Restrepo-Jimenez P, Pelak VS. Optical Coherence Tomography in Mild Cognitive Impairment: A Systematic Review and Meta-Analysis. *Front Neurol*. 2020;11:578698. doi:10.3389/fneur.2020.578698
16. Noah AM, Almghairbi D, Moppett IK. Optical coherence tomography in mild cognitive impairment - Systematic review and meta-analysis. *Clin Neurol Neurosurg*. 09 2020;196:106036. doi:10.1016/j.clineuro.2020.106036
17. Hui J, Zhao Y, Yu S, Liu J, Chiu K, Wang Y. Detection of retinal changes with optical coherence tomography angiography in mild cognitive impairment and Alzheimer's disease patients: A meta-analysis. *PLoS One*. 2021;16(8):e0255362. doi:10.1371/journal.pone.0255362
18. Jin Q, Lei Y, Wang R, Wu H, Ji K, Ling L. A Systematic Review and Meta-Analysis of Retinal Microvascular Features in Alzheimer's Disease. *Front Aging Neurosci*. 2021;13:683824. doi:10.3389/fnagi.2021.683824
19. Katsimpris A, Karamaounas A, Sideri AM, Katsimpris J, Georgalas I, Petrou P. Optical coherence tomography angiography in Alzheimer's disease: a systematic review and meta-analysis. *Eye (Lond)*. Jun 30 2021;doi:10.1038/s41433-021-01648-1

20. Rifai OM, McGrory S, Robbins CB, et al. The application of optical coherence tomography angiography in Alzheimer's disease: A systematic review. *Alzheimers Dement (Amst)*. 2021;13(1):e12149. doi:10.1002/dad2.12149
21. Zhang JF, Wiseman S, Valdés-Hernández MC, et al. The Application of Optical Coherence Tomography Angiography in Cerebral Small Vessel Disease, Ischemic Stroke, and Dementia: A Systematic Review. *Front Neurol*. 2020;11:1009. doi:10.3389/fneur.2020.01009
22. Lemmens S, Devulder A, Van Keer K, Bierkens J, De Boever P, Stalmans I. Systematic Review on Fractal Dimension of the Retinal Vasculature in Neurodegeneration and Stroke: Assessment of a Potential Biomarker. *Front Neurosci*. 2020;14:16. doi:10.3389/fnins.2020.00016
23. McGrory S, Cameron JR, Pellegrini E, et al. The application of retinal fundus camera imaging in dementia: A systematic review. *Alzheimers Dement (Amst)*. 2017;6:91-107. doi:10.1016/j.dadm.2016.11.001
24. Wu H, Wang C, Chen C, et al. Association between Retinal Vascular Geometric Changes and Cognitive Impairment: A Systematic Review and Meta-Analysis. *J Clin Neurol*. Jan 2020;16(1):19-28. doi:10.3988/jcn.2020.16.1.19
25. Snyder PJ, Johnson LN, Lim YY, et al. Nonvascular retinal imaging markers of preclinical Alzheimer's disease. *Alzheimers Dement (Amst)*. 2016;4:169-178. doi:10.1016/j.dadm.2016.09.001
26. Moja L, Fernandez del Rio MP, Banzi R, et al. Multiple systematic reviews: methods for assessing discordances of results. *Intern Emerg Med*. Dec 2012;7(6):563-8. doi:10.1007/s11739-012-0846-1
27. Lucenteforte E, Moja L, Pecoraro V, et al. Discordances originated by multiple meta-analyses on interventions for myocardial infarction: a systematic review. *J Clin Epidemiol*. Mar 2015;68(3):246-56. doi:10.1016/j.jclinepi.2014.11.004
28. Siontis KC, Hernandez-Boussard T, Ioannidis JP. Overlapping meta-analyses on the same topic: survey of published studies. *BMJ*. Jul 19 2013;347:f4501. doi:10.1136/bmj.f4501
29. Moher D, Tetzlaff J, Tricco AC, Sampson M, Altman DG. Epidemiology and reporting characteristics of systematic reviews. *PLoS Med*. Mar 27 2007;4(3):e78. doi:10.1371/journal.pmed.0040078
30. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *Rev Esp Cardiol (Engl Ed)*. Sep 2021;74(9):790-799. doi:10.1016/j.rec.2021.07.010
31. Whiting P, Savović J, Higgins JP, et al. ROBIS: A new tool to assess risk of bias in systematic reviews was developed. *J Clin Epidemiol*. Jan 2016;69:225-34. doi:10.1016/j.jclinepi.2015.06.005
32. Salgado JF. Transforming the Area under the Normal Curve (AUC) into Cohen's d, Pearson's r pb , Odds-Ratio, and Natural Log Odds-Ratio: Two Conversion Tables. The European Journal of Psychology Applied to Legal Context 2018. p. 35-47.
33. Mandrekar JN. Receiver operating characteristic curve in diagnostic test assessment. *J Thorac Oncol*. Sep 2010;5(9):1315-6. doi:10.1097/JTO.0b013e3181ec173d
34. Kim K, Yu SY, Kwak HW, Kim ES. Retinal Neurodegeneration Associated With Peripheral Nerve Conduction and Autonomic Nerve Function in Diabetic Patients. *Am J Ophthalmol*. Oct 2016;170:15-24. doi:10.1016/j.ajo.2016.06.038
35. Thomson KL, Yeo JM, Waddell B, Cameron JR, Pal S. A systematic review and meta-analysis of retinal nerve fiber layer change in dementia, using optical coherence tomography. *Alzheimers Dement (Amst)*. Jun 2015;1(2):136-43. doi:10.1016/j.dadm.2015.03.001
36. Coppola G, Di Renzo A, Ziccardi L, et al. Optical Coherence Tomography in Alzheimer's Disease: A Meta-Analysis. *PLoS One*. 2015;10(8):e0134750. doi:10.1371/journal.pone.0134750
37. Wang M, Zhu Y, Shi Z, Li C, Shen Y. Meta-analysis of the relationship of peripheral retinal nerve fiber layer thickness to Alzheimer's disease and mild cognitive impairment. *Shanghai Arch Psychiatry*. Oct 2015;27(5):263-79. doi:10.11919/j.issn.1002-0829.215100
38. Schönfeldt-Lecuona C, Schmidt A, Pinkhardt EH, et al. [Optical Coherence Tomography (OCT)--a new diagnostic tool in psychiatry?]. *Fortschr Neurol Psychiatr*. Oct 2014;82(10):566-71. doi:10.1055/s-0034-1385024
39. Ding J, Patton N, Deary IJ, et al. Retinal microvascular abnormalities and cognitive dysfunction: a systematic review. *Br J Ophthalmol*. Aug 2008;92(8):1017-25. doi:10.1136/bjo.2008.141994

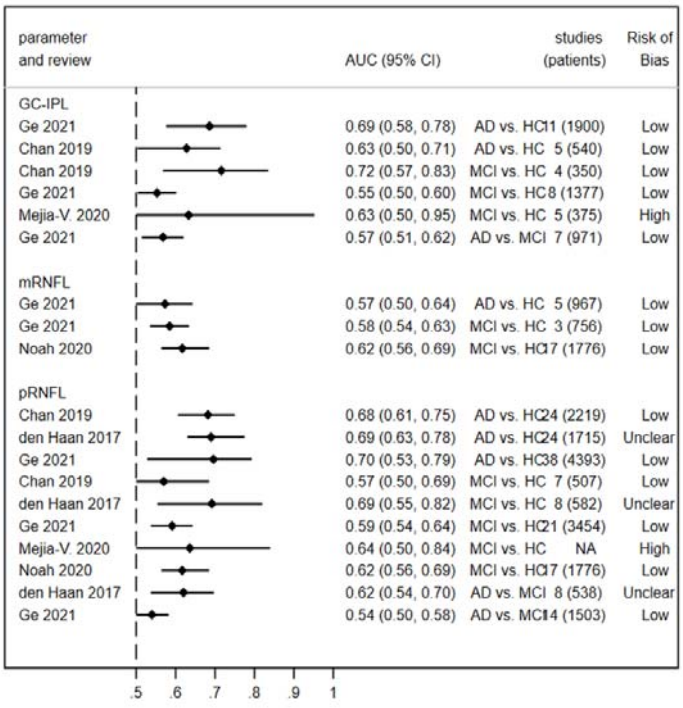
40. Heringa SM, Bouvy WH, van den Berg E, Moll AC, Kappelle LJ, Biessels GJ. Associations between retinal microvascular changes and dementia, cognitive functioning, and brain imaging abnormalities: a systematic review. *J Cereb Blood Flow Metab*. Jul 2013;33(7):983-95. doi:10.1038/jcbfm.2013.58
41. He XF, Liu YT, Peng C, Zhang F, Zhuang S, Zhang JS. Optical coherence tomography assessed retinal nerve fiber layer thickness in patients with Alzheimer's disease: a meta-analysis. *Int J Ophthalmol*. 2012;5(3):401-5. doi:10.3980/j.issn.2222-3959.2012.03.30
42. Kuźma E, Littlejohns TJ, Khawaja AP, Llewellyn DJ, Ukoumunne OC, Thiem U. Visual Impairment, Eye Diseases, and Dementia Risk: A Systematic Review and Meta-Analysis. *J Alzheimers Dis*. 2021;83(3):1073-1087. doi:10.3233/JAD-210250
43. Donahue RJ, Moller-Trane R, Nickells RW. Meta-analysis of transcriptomic changes in optic nerve injury and neurodegenerative models reveals a fundamental response to injury throughout the central nervous system. *Mol Vis*. 2017;23:987-1005.
44. Mardin CY, Hosari S. [Optical coherence tomography angiography in neuronal diseases : Preliminary findings]. *Ophthalmologe*. Aug 2019;116(8):714-721. doi:10.1007/s00347-019-0883-5
45. Youssef P, Nath S, Chaimowitz GA, Prat SS. Electroretinography in psychiatry: A systematic literature review. *Eur Psychiatry*. 10 2019;62:97-106. doi:10.1016/j.eurpsy.2019.09.006
46. Wu SZ, Masurkar AV, Balcer LJ. Afferent and Efferent Visual Markers of Alzheimer's Disease: A Review and Update in Early Stage Disease. *Front Aging Neurosci*. 2020;12:572337. doi:10.3389/fnagi.2020.572337
47. Opwonya J, Doan DNT, Kim SG, et al. Saccadic Eye Movement in Mild Cognitive Impairment and Alzheimer's Disease: A Systematic Review and Meta-Analysis. *Neuropsychol Rev*. May 06 2021;doi:10.1007/s11065-021-09495-3
48. Kahana Levy N, Lavidor M, Vakil E. Prosaccade and Antisaccade Paradigms in Persons with Alzheimer's Disease: A Meta-Analytic Review. *Neuropsychol Rev*. 03 2018;28(1):16-31. doi:10.1007/s11065-017-9362-4
49. Hadoux X, Hui F, Lim JKH, et al. Non-invasive in vivo hyperspectral imaging of the retina for potential biomarker use in Alzheimer's disease. *Nat Commun*. 09 17 2019;10(1):4227. doi:10.1038/s41467-019-12242-1
50. More SS, Beach JM, McClelland C, Mokhtarzadeh A, Vince R. In Vivo Assessment of Retinal Biomarkers by Hyperspectral Imaging: Early Detection of Alzheimer's Disease. *ACS Chem Neurosci*. 11 20 2019;10(11):4492-4501. doi:10.1021/acschemneuro.9b00331
51. Gopalakrishna G, Langendam MW, Scholten RJ, Bossuyt PM, Leeflang MM. Defining the clinical pathway in cochrane diagnostic test accuracy reviews. *BMC Med Res Methodol*. 11 10 2016;16(1):153. doi:10.1186/s12874-016-0252-x
52. Corvi F, Pellegrini M, Erba S, Cozzi M, Staurenghi G, Giani A. Reproducibility of Vessel Density, Fractal Dimension, and Foveal Avascular Zone Using 7 Different Optical Coherence Tomography Angiography Devices. *Am J Ophthalmol*. Feb 2018;186:25-31. doi:10.1016/j.ajo.2017.11.011
53. Thakoor KA, Li X, Tsamis E, Sajda P, Hood DC. Enhancing the Accuracy of Glaucoma Detection from OCT Probability Maps using Convolutional Neural Networks. *Annu Int Conf IEEE Eng Med Biol Soc*. Jul 2019;2019:2036-2040. doi:10.1109/EMBC.2019.8856899
54. Dong Y, Guo X, Arsiwala-Scheppach LT, et al. Association of Optical Coherence Tomography and Optical Coherence Tomography Angiography Retinal Features With Visual Function in Older Adults. *JAMA Ophthalmol*. Jul 14 2022;doi:10.1001/jamaophthalmol.2022.2099
55. Kashani AH. Promises and Pitfalls of Retinal Biomarkers in Systemic Health and Disease. *JAMA Ophthalmol*. Jul 14 2022;doi:10.1001/jamaophthalmol.2022.2100
56. Simel DL, Rennie D, Bossuyt PM. The STARD statement for reporting diagnostic accuracy studies: application to the history and physical examination. *J Gen Intern Med*. Jun 2008;23(6):768-74. doi:10.1007/s11606-008-0583-3

481 **Figure 1.**



482

483 **Figure 2.**



484

Figure 3.

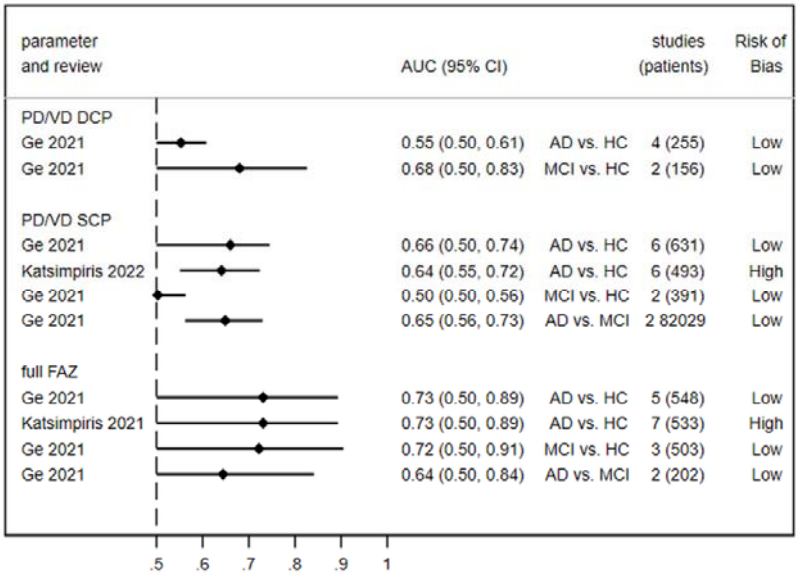


Figure 4.

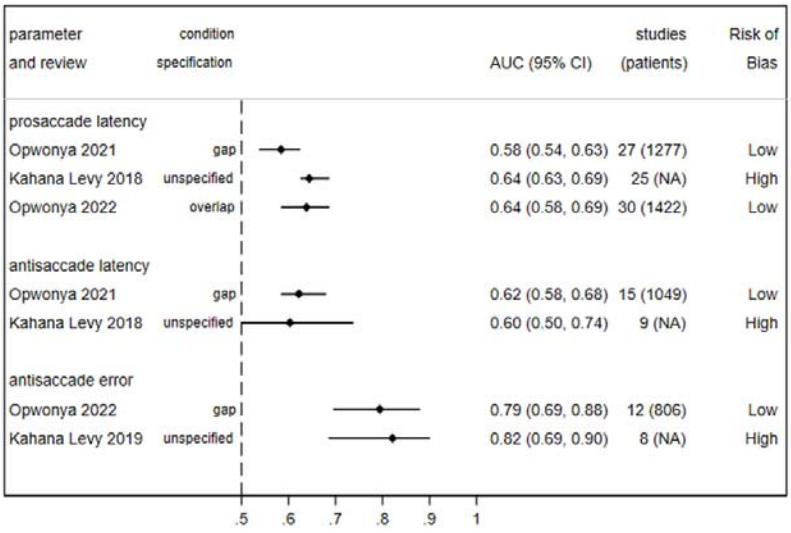


Figure legend

Figure 1.

Flow diagram of the literature search (PRISMA diagram)

Figure 2.

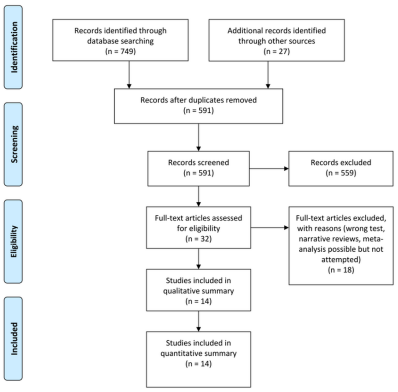
Forest plot of optical coherence tomography parameters for AD and MCI versus HC. All values are extensively reported in Supplementary file 3. Abbreviations: AD, Alzheimer's disease; MCI, Mild Cognitive Impairment; HC, health controls; AUC, Area Under the Curve; GC-IPL, ganglion cell-inner plexiform layer complex; mRNFL, macular retinal nerve fiber layer; pRNFL, peripapillary retinal nerve fiber layer.

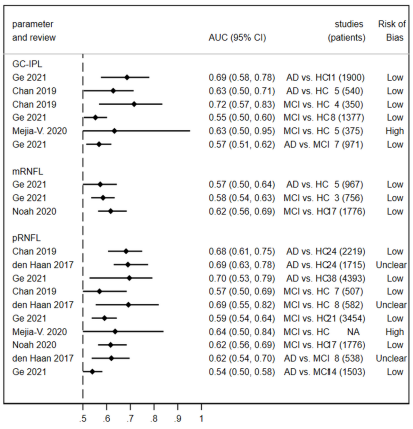
Figure 3.

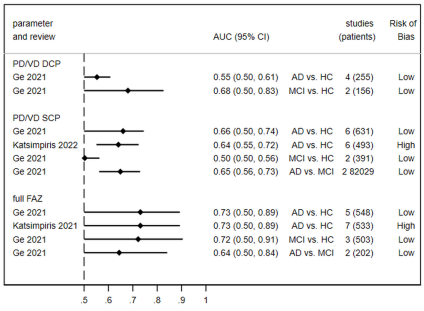
Forest plot of optical coherence tomography angiography parameters for AD and MCI versus HC. All values are extensively reported in Supplementary file 3. Abbreviations: AD, Alzheimer's disease; MCI, Mild Cognitive Impairment; HC, health controls; AUC, Area Under the Curve; PD, perfusion density; VD, vessel density; DCP, deep capillary plexus; SCP, superficial capillary plexus; FAZ, foveal avascular zone.

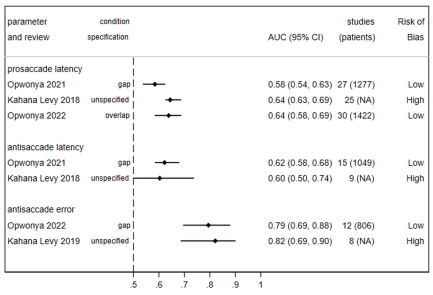
Figure 4.

Forest plot of saccadic eye movements' parameters for AD and MCI versus HC. All values are extensively reported in eTable 3 in the Supplement.


Identification
Screening
Eligibility
Included
Records identified through database searching (n = 749)
Additional records identified through other sources (n = 27)
Records after duplicates removed (n = 591)
Records screened (n = 591)
Records excluded (n = 559)
Full-text articles assessed for eligibility (n = 32)
Full-text articles excluded, with reasons (wrong test, narrative reviews, meta-analysis possible but not attempted) (n = 18)
Studies included in qualitative summary (n = 14)
Studies included in quantitative summary (n = 14)


parameter and review
AUC (95% CI)
studies (patients)
Risk of Bias
GC-IPL
Ge 2021 0.69 (0.58, 0.78) AD vs. HC 11 (1900) Low
Chan 2019 0.63 (0.50, 0.71) AD vs. HC 5 (540) Low
Chan 2019 0.72 (0.57, 0.83) MCI vs. HC 4 (350) Low
Ge 2021 0.55 (0.50, 0.60) MCI vs. HC 8 (1377) Low
Mejia-V. 2020 0.63 (0.50, 0.95) MCI vs. HC 5 (375) High
Ge 2021 0.57 (0.51, 0.62) AD vs. MCI 7 (971) Low
mRNFL
Ge 2021 0.57 (0.50, 0.64) AD vs. HC 5 (967) Low
Ge 2021 0.58 (0.54, 0.63) MCI vs. HC 3 (756) Low
Noah 2020 0.62 (0.56, 0.69) MCI vs. HC 7 (1776) Low
pRNFL
Chan 2019 0.68 (0.61, 0.75) AD vs. HC 24 (2219) Low
den Haan 2017 0.69 (0.63, 0.78) AD vs. HC 24 (1715) Unclear
Ge 2021 0.70 (0.53, 0.79) AD vs. HC 38 (4393) Low
Chan 2019 0.57 (0.50, 0.69) MCI vs. HC 7 (507) Low
den Haan 2017 0.69 (0.55, 0.82) MCI vs. HC 8 (582) Unclear
Ge 2021 0.59 (0.54, 0.64) MCI vs. HC 21 (3454) Low
Mejia-V. 2020 0.64 (0.50, 0.84) MCI vs. HC NA High
Noah 2020 0.62 (0.56, 0.69) MCI vs. HC 7 (1776) Low
den Haan 2017 0.62 (0.54, 0.70) AD vs. MCI 8 (538) Unclear
Ge 2021 0.54 (0.50, 0.58) AD vs. MCI 4 (1503) Low
.5
.6
.7
.8
.9
1


parameter and review
AUC (95% CI)
studies (patients)
Risk of Bias
PD/VD DCP
Ge 2021 0.55 (0.50, 0.61) AD vs. HC 4 (255) Low
Ge 2021 0.68 (0.50, 0.83) MCI vs. HC 2 (156) Low
PD/VD SCP
Ge 2021 0.66 (0.50, 0.74) AD vs. HC 6 (631) Low
Katsimpiris 2022 0.64 (0.55, 0.72) AD vs. HC 6 (493) High
Ge 2021 0.50 (0.50, 0.56) MCI vs. HC 2 (391) Low
Ge 2021 0.65 (0.56, 0.73) AD vs. MCI 2 82029 Low
full FAZ
Ge 2021 0.73 (0.50, 0.89) AD vs. HC 5 (548) Low
Katsimpiris 2021 0.73 (0.50, 0.89) AD vs. HC 7 (533) High
Ge 2021 0.72 (0.50, 0.91) MCI vs. HC 3 (503) Low
Ge 2021 0.64 (0.50, 0.84) AD vs. MCI 2 (202) Low
.5
.6
.7
.8
.9
1


parameter and review
condition specification
AUC (95% CI)
studies (patients)
Risk of Bias
prosaccade latency
Opwonya 2021 gap 0.58 (0.54, 0.63) 27 (1277) Low
Kahana Levy 2018 unspecified 0.64 (0.63, 0.69) 25 (NA) High
Opwonya 2022 overlap 0.64 (0.58, 0.69) 30 (1422) Low
antisaccade latency
Opwonya 2021 gap 0.62 (0.58, 0.68) 15 (1049) Low
Kahana Levy 2018 unspecified 0.60 (0.50, 0.74) 9 (NA) High
antisaccade error
Opwonya 2022 gap 0.79 (0.69, 0.88) 12 (806) Low
Kahana Levy 2019 unspecified 0.82 (0.69, 0.90) 8 (NA) High
.5
.6
.7
.8
.9
1