

Review

Detection and Management of Geographic Atrophy Secondary to Age-Related Macular Degeneration Using Noninvasive Retinal Images and Artificial Intelligence: Systematic Review

Nannan Shi^{1*}, MD, PhD; Jiaxian Li^{2*}, MD, PhD; Mengqiu Shang¹, MD, PhD; Weidao Zhang¹, MD, PhD; Kai Xu¹, MD, PhD; Yamin Li¹, MD, PhD; Lina Liang¹, MD, PhD

¹Department of Eye Function Laboratory, Eye Hospital China Academy of Chinese Medical Sciences, Beijing, China

²Department of Ophthalmology, The First Affiliated Hospital of Yunnan University of Chinese Medicine, Yunnan Provincial Hospital of Traditional Chinese Medicine, Kunming, China

*these authors contributed equally

Corresponding Author:

Lina Liang, MD, PhD
Department of Eye Function Laboratory
Eye Hospital China Academy of Chinese Medical Sciences
No.33 Lugu Road, Shijingshan District
Beijing 100040
China
Phone: +86 010-68683451
Email: lianglina163@163.com

Abstract

Background: Geographic atrophy (GA), the endpoint of dry age-related macular degeneration (AMD), is irreversible. The recent approval by the Food and Drug Administration of a complement component 3 inhibitor marks a significant breakthrough, highlighting the critical importance of early detection and management of GA. Consequently, there is an urgent and unmet need for efficient, accurate, and accessible methods to identify and monitor GA. Artificial intelligence (AI), particularly deep learning (DL), applied to noninvasive retinal imaging, offers a promising solution for automating and enhancing GA management.

Objective: This systematic review aimed to assess the performance of AI using noninvasive imaging modalities and compare it with clinical expert assessment as the ground truth.

Methods: Two consecutive searches were conducted on PubMed, Embase, Web of Science, Scopus, Cochrane Library, and CINAHL. The last search was performed on October 5, 2025. Studies using AI for GA secondary to dry AMD via noninvasive retinal imaging were included. Two authors worked in pairs to extract the study characteristics independently. A third author adjudicated disagreements. Quality Assessment of Diagnostic Accuracy Studies-AI and Prediction Model Risk of Bias Assessment Tool (PROBAST) were applied to evaluate the risk of bias and application.

Results: Of the 803 records initially identified, 176 were found through an updated search. Subsequently, 200 papers were assessed in full text, of which 41 were included in the final analysis, 10 for GA detection, 20 for GA assessment and progression, and 11 for GA lesion prediction. The reviewed studies collectively involved at least 24,592 participants (detection: n=7132, assessment and progression: n=14,064, and prediction: n=6706), with a wide age range of 50 to 94 years. The studies spanned a diverse array of countries, including the United States, the United Kingdom, China, Austria, Australia, France, Israel, Italy, Switzerland, and Germany, as well as a multicenter study encompassing 7 European nations. The studies used a variety of imaging modalities to assess GA, including color fundus photography, fundus autofluorescence, near-infrared reflectance, spectral domain-optical coherence tomography (OCT), swept-source (SS)-OCT, and 3D-OCT. DL algorithms (eg, U-Net, ResNet50, EfficientNetB4, Xception, Inception v3, and PSC-UNet) consistently showed remarkable performance in GA detection and management tasks, with several studies achieving performance comparable to clinical experts.

Conclusions: AI, particularly DL-based algorithms, holds considerable promise for the detection and management of GA secondary to dry AMD with performance comparable to ophthalmologists. This review innovatively consolidates evidence

across GA management—from initial detection to progression prediction—using diverse noninvasive imaging. It has strong potential to augment clinical decision-making. However, to realize this potential in real-world settings, future research is needed to robustly enhance reporting specifications, ensure data diversity across populations and devices, and implement rigorous external validation in prospective, multicenter studies.

Trial Registration: PROSPERO CRD420251000963; <https://www.crd.york.ac.uk/PROSPERO/view/CRD420251000963>

J Med Internet Res 2025;27:e81328; doi: [10.2196/81328](https://doi.org/10.2196/81328)

Keywords: geographic atrophy; dry age-related macular degeneration; artificial intelligence; noninvasive retinal images; systematic review

Introduction

Age-related macular degeneration (AMD) is a progressive retinal disorder affecting millions of people worldwide [1]. In its advanced stages, characterized by neovascularization and geographic atrophy (GA), it can lead to significant vision loss, although symptoms may be subtle during the early and intermediate phases [2]. The Classification of Atrophy Meetings group has defined atrophy lesion development as incomplete retinal pigment epithelium (RPE) and outer retinal atrophy and complete RPE and outer retinal atrophy (cRORA) based on imaging methods [3]. GA, also known as cRORA, is the endpoint of dry AMD and is characterized by the loss of photoreceptors, RPE, and choriocapillaris [4,5]. With the advent of 2 approved therapies for GA secondary to AMD in 2023, namely pegcetacoplan (Syfovre) [6] and avacincaptad pegol [7], the treatment of GA represents a significant breakthrough. However, the effectiveness of these therapies relies heavily on early detection and the ability to monitor treatment response—a significant unmet need in current clinical practice. The recent approval of complement inhibitors underscores the necessity for precise, reproducible, and practical tools to not only identify GA at its earliest stages but also to objectively track morphological changes over time, thereby evaluating therapeutic efficacy [8,9]. Artificial intelligence (AI) is uniquely positioned to address this gap by enabling precise, reproducible, and automated quantification of GA progression and treatment response using noninvasive imaging modalities [10]. Unlike conventional methods that rely on subjective and time-consuming manual assessments, AI algorithms can detect subtle subclinical changes in retinal structures—such as photoreceptor integrity loss, RPE atrophy, and hyperreflective foci—long before they become clinically apparent. Thus, AI-based retinal imaging offers a critical foundation for early detection and timely intervention in GA.

Various imaging techniques, both invasive and noninvasive, can directly visualize GA lesions. Invasive methods, such as fluorescence angiography, often result in a poor patient experience and entail high costs due to pupil dilation and sodium fluorescein injection. While it remains the gold standard for assessing neovascular AMD and offers significant diagnostic insights for retinal vascular diseases, in most cases, noninvasive fundus images are used for GA diagnosis and management [2]. Color fundus photography (CFP), fundus autofluorescence (FAF), and near-infrared reflectance (NIR) are based on 2D images, which can generally produce

results to quantify the atrophic area but fail to identify the retinal structure axially [11]. Compared with fundus imaging, optical coherence tomography (OCT) provides high-resolution, noninvasive 3D images of retinal structures for macular assessment. In addition, conventional B-scan (axial direction) OCT images can be integrated with en-face scans, facilitating the identification of atrophy borders similar to FAF [10,12]. Nonetheless, manual labeling is tedious, time-consuming, and impractical in a clinical setup [13]. There is an urgent and unmet need for early detection and management of GA using retinal image modalities. Recent advancements in AI, especially deep learning (DL), present a promising opportunity for enhancing GA detection, classification, segmentation, quantification, and prediction.

In the 1950s, AI referred to computer systems capable of performing complex tasks that historically only a human could do. So what is AI? How is it used in medicine today? And what may it do in the future? AI refers to the theory and development of computer systems capable of performing tasks that historically required human intelligence, such as recognizing speech, making decisions, and identifying patterns. AI is an umbrella term that encompasses a wide variety of technologies, including machine learning (ML) and DL [14]. ML is a subfield of AI that uses algorithms trained on datasets to create self-learning models capable of predicting outcomes and classifying information without human intervention [15]. ML refers to the general use of algorithms and data to create autonomous or semiautonomous machines. DL, meanwhile, is a subset of ML that layers algorithms into “neural networks” with 3 or more layers. Thus, it somewhat resembles the human brain, enabling machines to perform increasingly complex tasks [16]. DL algorithms generally have high and clinically acceptable diagnostic accuracy across different areas (ophthalmology, respiratory, breast cancer, etc) in radiology [17]. Within ophthalmology, DL algorithms showed reliable performance for detecting multiple findings in macular-centered retinal fundus images [18]. Therefore, automatic GA segmentation plays a vital role in the diagnosis and management of advanced AMD and its application in the clinical setting.

Given the rapid evolution of AI applications in ophthalmology and the growing clinical importance of GA, this study aimed to systematically review the current evidence on AI-based approaches for the detection and management of GA secondary to dry AMD using noninvasive imaging modalities. We aimed to evaluate diagnostic accuracy relative to reference standards and examine methodological

challenges to inform the design of future research and clinical implementation.

Methods

Protocol and Registration

Before starting this systematic review and meta-analysis, we registered a protocol on the PROSPERO website. This review adhered to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) and PRISMA-DTA (PRISMA of Diagnostic Test Accuracy) checklists [19, 20].

Eligibility Criteria

We included studies using AI algorithms to detect, classify, identify, segment, quantify, or predict GA secondary to AMD from CFP, OCT, OCT angiography, FAF, or NIR. The data were from participants, with or without symptoms, who were diagnosed with GA (or cRORA) secondary to nonexudative AMD. Study designs were not restricted; multicenter or single-center, prospective or retrospective, post hoc analysis, clinical study, or model development studies were all accepted. Eyes with neovascular complications or macular atrophy from causes other than AMD, any previous anti-vascular endothelial growth factor treatment, any confounding retinopathy, or poor image quality were excluded.

Electronic Search Strategy

Two consecutive searches were conducted on PubMed, Embase, Web of Science, Scopus, Cochrane Library, and CINAHL. Because this review required the extraction of baseline data and items, considering the completeness of the data, we did not conduct any in press or print source searches and excluded conference proceedings and similar materials. The initial search was completed from the date of entry to December 1, 2024; the updated search, from December 1, 2024, to October 5, 2025. We used a search strategy for patient (GA) and index tests (AI and retinal images) that had been used in previous Cochrane Review without any search peer review process [21]. There were no restrictions on the date of publication. The language was limited to English. In [Multimedia Appendix 1](#), detailed search strategies for each database are provided. During this process, no filters were used. During the search process, we adhered to the PRISMA-S (Preferred Reporting Items for Systematic reviews and Meta-Analyses literature search extension) reporting guidelines [22].

Selection Process

All relevant literature was imported into EndNote (version 20; Clarivate Analytics) software, and literature screening was conducted independently by 2 researchers (NS and JL) who specialize in ophthalmology. Duplicates were removed from the software, and the titles and abstracts of the literature were reviewed to identify those relevant to the topic. Finally, the full texts were downloaded and examined, leading to the selection of literature that met the inclusion criteria. In cases of inconsistencies in the final inclusion decisions made by

the 2 researchers, a third professional (LL) was consulted to resolve the dispute.

Data Collection Process

Using standardized data items, the data were extracted independently from the included studies by 2 researchers (NS and JL). A third review author (LL) confirmed or adjudicated any discrepancies through group discussion. We retrieved the following data items: (1) study characteristics (author, year, study design, region, and theme), (2) dataset characteristics (databases, source of databases, training/validation/testing ratio, patient number, number of images or volumes, scan number, mean age, clinical registration number, and model evaluation method), (3) image and algorithm characteristics (devices, metrics, image modality, image resolution, and AI algorithms), (4) performance metrics (outcomes, performance of models, ground truth, and performance of the ophthalmologists), and (5) main results. All the information was retrieved from the main text and the tables provided in [Multimedia Appendix 2](#). Therefore, we did not seek additional data by contacting the authors or experts. In some studies, the authors reported multiple sets of performance data based on a subset of a single dataset. For example, they may have reported results such as sensitivity, specificity, accuracy, and so forth, conducted on the cross-validation set, the test set, or the development set. We referred to the relevant literature to select the optimal set of test performance results [21]. However, when the primary study provided performance results based on a single test, the development dataset was used to train the AI model, and an external validation set ultimately was used to determine the performance of the optimal model. We extracted the external validation set performance data [23].

Risk of Bias and Application

We worked in pairs to assess the risk of bias and the applicability of the studies, which involved detection, classification, identification, segmentation, and quantification using the Quality Assessment of Diagnostic Accuracy Studies (QUADAS)-AI [24] and the modified QUADAS-2 tool [25], while predictive studies used the Prediction Model Risk of Bias Assessment Tool (PROBAST) [26].

In the current context, QUADAS-AI has not yet established a complete specification of items. Therefore, we referenced the examples provided by QUASAS-AI and the published literature to compile the revised QUADAS-AI items, which included 4 domains and 9 leading questions (Table S4 in [Multimedia Appendix 2](#)). The PROBAST tool comprises participants, predictors, outcomes, and analysis, containing 20 signaling questions across 4 domains (Table S5 in [Multimedia Appendix 2](#)). We also evaluated the applicability of the study based on the leading or signaling questions in the first 3 domains. A study with “yes” answers to all index questions was considered to have a low risk of bias. If the answer to any of the informational questions was “no,” there was a potential for bias, leading the authors to rate the risk of bias as high. “Indeterminate” grades were applied when detailed content was not provided in the literature, making it difficult for the evaluator to reach a judgment. They were

used only when the reported data were insufficient. Throughout the process, disagreements between the 2 reviewers (NS and JL) were resolved by consulting the senior reviewer (LL).

Data Synthesis

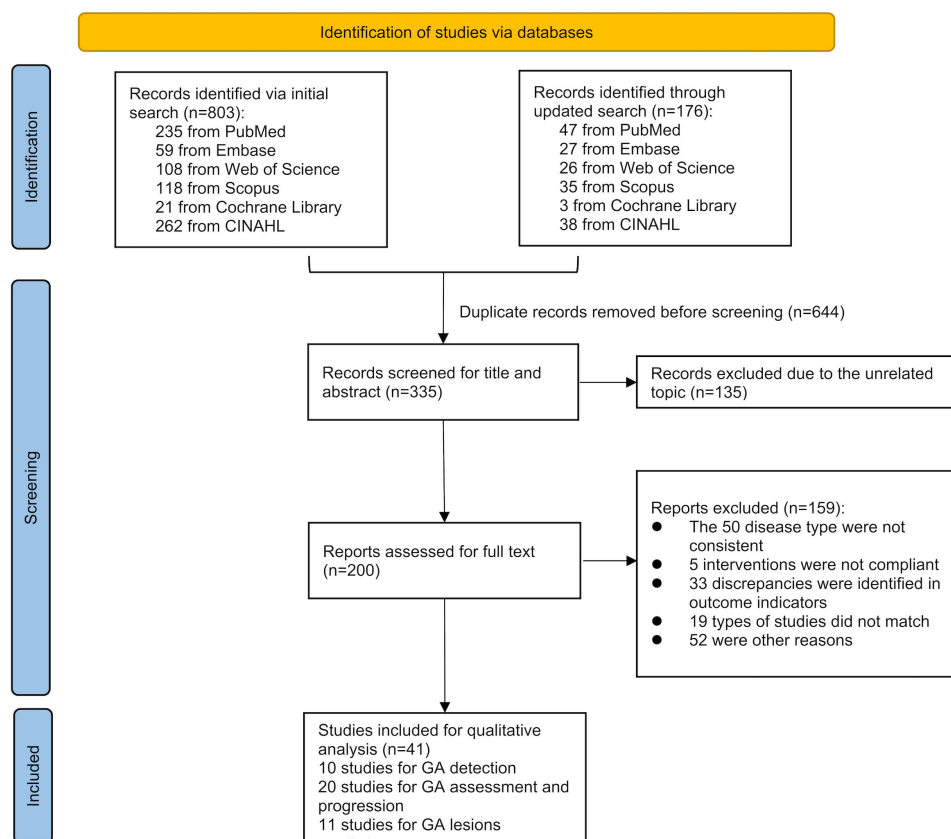
As very few studies reported the number of true positives, true negatives, false positives, and false negatives, we restricted the quantitative analysis to determine the diagnostic accuracy of AI as a triaging tool for GA secondary to nonexudative AMD. However, a meta-analysis was not performed due to significant methodological heterogeneity across studies, arising from diverse AI architectures, imaging modalities, outcome metrics, and validation protocols. Instead, a systematic review was performed to qualitatively summarize performance trends. This approach allowed for a comprehensive evaluation of the AI capabilities in the detection and management of GA via noninvasive images.

Results

Study Selection

A total of 979 records related to the topic of this systematic review were searched across 6 different databases using a combination of subject terms and free-text terms. After removing duplicates, 335 records remained and were examined for titles and abstracts. Excluding studies not relevant to the research topic resulted in 200 reports. The full texts were then downloaded and reviewed in detail based on the eligibility criteria for the studies. In the final qualitative analysis, 41 studies were included. Of these, 10 focusing on GA diagnosis, 20 on GA assessment and progression, and 11 on GA prediction. [Figure 1](#) presents the detailed flow diagram of the literature selection.

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram for literature selection. GA: geographic atrophy.



AI in Detecting the Presence of GA

Ten of the 41 included studies focused on AI-based detection of GA using noninvasive retinal images (Table S1 in [Multimedia Appendix 2](#)). As listed in [Table 1](#), the studies were published from 2018 to 2025. Four of the studies [27-30] focused on model development, 3 [31-33] were retrospective studies, and 3 [34-36] were prospective studies (1 multicenter cohort study, 1 multicenter and low-interventional clinical study, and 1 clinical study). Geographically, half were from the United States, with others from Israel, Italy,

Switzerland, Germany, and a multicenter European collaboration. The studies addressed several detection-related tasks: 5 focused solely on GA detection [27-29,34,35], 2 covered detection and classification [30,33], and others integrated detection with quantification or segmentation [31,32,36].

Table 1. Characteristics of studies evaluating artificial intelligence (AI) models for geographic atrophy (GA) detection using noninvasive retinal imaging.

Author	Study design	Region	Purpose of the study	Source of datasets	Number of patients	Number of images or scans	Model evaluation method	Image modality (image resolution)	AI ^a algorithms	Outcomes	Performance of models
Fineberg et al [33]	Retrospective cohort study	Israel (Petah Tikva)	Detection and classification (GA)	Rabin Medical Center	113	659	10-fold cross-validation	NIR (640*640 pixels)	CNNs: ^d ResNet50, EfficientNetB0, ViT_B_16, and YOLOv8 variants.	ACC, P, SEN, SPE, F ₁ , IoU ^j , and DSC ^g	<ul style="list-style-type: none">GA classification: EfficientNetB0: ACC^b=0.9148; P^m=0.9204; SEN^p=0.9233; SPE^q=1.0; F₁=0.9147.ResNet50: ACC=0.8815; P=.8933; SEN=0.8917; SPE=0.9833; F₁=0.8812.ViT_B_16: ACC=0.963; P=.9632; SEN=0.9667; SPE=1.0; F₁=0.9629.GA detection: YOLOv8-Large: SEN=0.91; P=0.91; IoU=0.84; DSC^e=0.88.
Kalra et al [31]	Retrospective clinical study	United States (Cleveland)	Detection, quantification, and segmentation (presence of GA and pixel-wise GA area measurement)	the Cole Eye Institute of the Cleveland Clinic	341	900	triple-fold cross-validation	SD-OCT ^o (256*256 pixels)	CNN ^d : U-Net	F ₁ , ACC, P, R ⁿ , SEN, and SPE	<ul style="list-style-type: none">GA detection: ACC=0.91, SEN=0.86, SPE=0.94, F₁=0.87.GA segmentation:

Author	Study design	Region	Purpose of the study	Source of datasets	Number of patients	Number of images or scans	Model evaluation method	Image modality (image resolution)	AI ^a algorithms	Outcomes	Performance of models
Derradji et al [32]	Retrospective clinical study	Switzerland (Lausanne)	Detection and segmentation (RORA)	An existing image database of the Medical Retina Department at Jules-Gonin Eye Hospital	57	62	5-fold cross-validation	SD-OCT (NR)	CNN: U-Net	SEN, DSC, <i>P</i> , and Kappa	ACC=0.96, SEN=0.95, SPE=0.93, <i>F</i> ₁ =0.82. • Grader 1: DSC: mean 0.881 (SD 0.074); Precision: mean 0.928 (SD 0.054); SEN: mean 0.850 (SD 0.119); Kappa: mean 0.846 (SD 0.072). • Grader 2: DSC: mean 0.844 (SD 0.076); Precision: mean 0.799 (SD 0.133); SEN: mean 0.915 (SD 0.064); Kappa: mean 0.800 (SD 0.082). • ZEISS: SEN=0.6; SPE=0.964; PPV=0.375; NPV=0.985.
de Vente et al [36]	Prospective multicenter and low-interventional clinical study (including cross-sectional and	20 sites in 7 European countries	Detection and quantification (cRORA ^f)	The MACUSTAR Study Cohort	168	143 (ZEISS); 167 (Spectralis)	NR	SD-OCT (512*650 pixels)	CNN: U-Net	SEN, SPE, PPV ^l , NPV ^l , and Kappa	

Author	Study design	Region	Purpose of the study	Source of datasets	Number of patients	Number of images or scans	Model evaluation method	Image modality (image resolution)	AI ^a algorithms	Outcomes	Performance of models
	longitudinal study part)										<ul style="list-style-type: none">Spectralis: SEN=0.625; SPE=0.974; PPV=0.714; NPV=0.961.SEN: 100% (95%CI 83.2%-100%); SPE=97.5% (95% CI 86.8%-99.9%) ; ACC=98.4%; F₁=0.976; R=1; AUROC^r=0.988 (95% CI 0.918-1); AUPRC^c=0.952 (95%CI 0.719-0.994).ACC=0.965 (95% CI 0.959-0.971); Kappa=0.611 (95% CI 0.533-0.689); SEN=0.692 (95% CI 0.560-0.825); SPE=0.978 (95% CI 0.970-0.985); Precision=0.584 (95% CI 0.491-0.676).
Sarao et al [35]	Prospective clinical study	Italy (Udine)	Detection (presence of GA)	the Istituto Europeo di Microchirurgia Oculare (IEMO) study	180	540	NR	CFP ^o (NR)	CNN: Efficientnet_b2	SEN, SPE, ACC, F ₁ , R, AUROC, and AUPRC ^c	
Keenan et al [34]	Multicenter and prospective cohort study	United States (Maryland)	Detection (presence of GA)	Age-Related Eye Disease Study (AREDS) dataset	4582	59,812	5-fold cross-validation	CFP (512 pixels)	CNN: inception v3	ACC, SEN, SPE, P, AUC, and Kappa	

Author	Study design	Region	Purpose of the study	Source of datasets	Number of patients	Number of images or scans	Model evaluation method	Image modality (image resolution)	AI ^a algorithms	Outcomes	Performance of models
Yao et al [29]	Model development and evaluation	United States (California)	Detection (presence of nGA)	the Early Stages of AMD ^b (LEAD) study	140	1884	5-fold cross-validation	SD-OCT (512*496 pixels)	CNN: ResNet18	SEN, SPE, ACC, P, and F ₁	• SEN=0.76 (95% CI 0.67-0.84); SPE=0.98 (95% CI 0.96-0.99); PRE=0.73 (95% CI 0.54-0.89); ACC=0.97 (95% CI 0.95-0.98); F ₁ =0.74 (95% CI 0.61-0.84). • SEN=0.909 (95% CI 0.778-1.000); SPE=0.553 (95% CI 0.394-0.703); PPV=0.541 (95% CI 0.375-0.707); NPV=0.913 (95% CI 0.778-1.000); AUROC=0.84 (95% CI 0.75-0.94); AUPRC=0.82 (95% CI 0.70-0.93). • ACC=0.93 (95% CI 0.92-0.94); Precision=0.9
Chiang et al [27]	Model development	United States (California)	Detection (complete retinal pigment epithelial and outer retinal atrophy (cRORA) in eyes with AMD)	University of Pennsylvania, University of Miami, and Case Western Reserve University; (2) Doheny Image Reading Research Laboratory, Doheny-UCLA (University of California Los Angeles Eye Centers)	71 (training); 649 (testing #1); 60 (testing #2)	188 (training); 1117 (testing #1)	5-fold cross-validation	SD-OCT (256*256 pixels)	CNN: ResNet18	SEN, SPE, PPV, NPV, AUROC, and AUPRC	• SEN=0.909 (95% CI 0.778-1.000); SPE=0.553 (95% CI 0.394-0.703); PPV=0.541 (95% CI 0.375-0.707); NPV=0.913 (95% CI 0.778-1.000); AUROC=0.84 (95% CI 0.75-0.94); AUPRC=0.82 (95% CI 0.70-0.93). • ACC=0.93 (95% CI 0.92-0.94); Precision=0.9
Elsawy et al [28]	Model development	United States (Maryland)	Detection (explain decision making and compare methods)	The Age-Related Eye Disease Study 2 (AREDS2)	311	1284 scans	10-fold cross-validation	SD-OCT (128*128 or 224* pixels)	3D CNN: deep-GA-Net	ACC, P, R, F ₁ , Kappa, AUROC, and AUPRC	• ACC=0.93 (95% CI 0.92-0.94); Precision=0.9

Author	Study design	Region	Purpose of the study	Source of datasets	Number of patients	Number of images or scans	Model evaluation method	Image modality (image resolution)	AI ^a algorithms	Outcomes	Performance of models
Treder et al [30]	Model development	Germany (Muenster)	Detection and classification (GA)	Public database: ImageNet	400 (training); 60 (test set)	400 (training); 60 (test set)	NR	FAF ^b (NR)	Deep CNN: self-learning algorithm	SEN, SPE, and ACC	0 (95% CI 0.88-0.91); Recall=0.90 (95% CI 0.89-0.92); <i>F</i> ₁ score=0.90 (95% CI 0.89-0.91); Kappa=0.80 (95% CI 0.77-0.83); AUROC=0.94 (95% CI 0.93-0.95); AUPRC=0.91 (95% CI 0.90-0.93). <ul style="list-style-type: none">• Probability score: mean 0.981 (SD 0.048); SEN=100%;

Author	Study design	Region	Purpose of the study	Source of datasets	Number of patients	Number of images or scans	Model evaluation method	Image modality (image resolution)	AI ^a algorithms	Outcomes	Performance of models
											SPE=100%; ACC=100%.
	^a AI: artificial intelligence.										
	^b ACC: accuracy.										
	^c AUPRC: area under the precision-recall curve.										
	^d CNN: convolutional neural network.										
	^e CFP: color fundus photography.										
	^f cRORA: complete retinal pigment epithelium and outer retinal atrophy.										
	^g DSC: dice similarity coefficient.										
	^h FAF: fundus autofluorescence.										
	ⁱ IoU: intersection over union.										
	^j NR: not reported.										
	^k OCT: optical coherence tomography.										
	^l PPV: positive predictive value.										
	^m P: precision.										
	ⁿ R: recall.										
	^o SD-OCT: spectral domain OCT.										
	^p SEN: sensitivity.										
	^q SPE: specificity.										
	^r AUROC: area under the receiver operating characteristic curve.										
	^s AMD: age-related macular degeneration.										
	^t NPV: negative predictive value.										

Dataset configurations varied: 6 studies used training, validation, and test sets [30-33,35,36]; 3 used only training and test sets [27,28,34]; and 1 included a tuning set [29]. Collectively, these studies involved at least 7132 participants, with ages ranging from 50 to 85 years. Three studies were registered with ClinicalTrials.gov (NCT00734487, NCT01790802, and NCT03349801) [28,29,36]. Cross-validation methods included 5-fold (40% of studies) [27,29,32,34], 10-fold (20%) [28,33], and triple-fold (10%) [31]; 30% did not report validation details.

Spectral-domain (SD)-OCT was the most frequently used imaging modality (6/10 of studies) [27-29,31,32,36], followed by CFP (2/10) [34,35], and FAF or NIR (2/10 each) [30,33]. Most studies applied image preprocessing techniques—such as size standardization, orientation adjustment, intensity normalization, and noise reduction—to improve model performance. DL-based algorithms for GA detection have been developed for multiple image modalities. For example, Derradji et al [32] trained a convolutional neural networks (CNNs), U-Net, to predict atrophic signs in the retina, based on the EfficientNet-b3 architecture. Kalra et al [31] and de Vente et al [36] also trained a DL model based on U-Net. Yao et al [29] applied 3D OCT scans with ResNet18 pretrained on the ImageNet dataset, and Chiang et al [27] developed CNN (ResNet18) to improve computational efficiency. Elsayw et al [28] proposed Deep-GA-Net, a 3D backbone CNN with a 3D loss-based attention layer, and evaluated the effectiveness of using attention layers. Sarao et al [35] used a deep CNN, the EfficientNet_b2 model, which was pretrained on the ImageNet dataset and is well-known for its high efficiency and small size. Keenan et al [34] established their model using Inception v3, while Treder et al [30] performed a deep CNN, a self-learning algorithm, processing input data with FAF images.

A total of 14 performance sets were extracted from the 10 studies. Key metrics included sensitivity, specificity, accuracy, positive predictive value, negative predictive value, intersection over union, area under the receiver operating characteristic curve, area under the precision-recall curve, F_1 -score, precision, recall, Kappa, and dice similarity coefficient. Six OCT-based studies showed that DL models could detect GA with high accuracy, comparable to human graders [27-29,31,32,36]. Two studies using CFP also reported strong performance [34,35], while FAF- and NIR-based approaches demonstrated excellent repeatability and reliability [30,33].

We conducted a thorough evaluation of the 10 diagnostic studies' methodological quality for the "participant selection," "index test," "reference standard," and "flow and timing" domains at the study level (Table 2). None of the studies had an overall low or unclear risk of bias; instead, every study had a high risk of bias in at least 1 of the 4 domains. Regarding "patient selection," only 4 studies [29,32,33,36] described the eligibility criteria; the rest did not report them. One study [30] used an open dataset (ImageNet) and did not include a test set. The small sample size of 4 studies [32,33,35,36] may have resulted in overfitting. In addition, 3 studies [29,30,32] did not report image formats and resolutions. Five studies [30,31,34-36] had a high risk of bias in participant selection because the included participants were not only GA secondary to dry AMD but also had other unrelated diseases. Regarding the "Index test," only 1 algorithm was externally validated using a different dataset [27]; all other items were evaluated as low risk.

Table 2. Methodological quality and applicability assessment for studies on geographic atrophy (GA) detection using the revised Quality Assessment of Diagnostic Accuracy Studies–Artificial Intelligence (QUADAS-AI).

Study	Risk of bias				Concerns regarding applicability		
	Patient selection	Index test	Reference standard	Flow and timing	Patient selection	Index test	Reference standard
Chiang et al [27]	High risk	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk
Elsawy et al [28]	High risk	High risk	Low risk	Low risk	Low risk	Low risk	Low risk
Kalra et al [31]	High risk	High risk	Low risk	Low risk	High risk	Low risk	Low risk
Keenan et al [34]	High risk	High risk	Low risk	Low risk	High risk	Low risk	Low risk
Sarao et al [35]	High risk	High risk	Low risk	Low risk	High risk	Low risk	Low risk
Yao et al [29]	High risk	High risk	Low risk	Low risk	Low risk	Low risk	Low risk
Treder et al [30]	High risk	High risk	Low risk	Low risk	High risk	Low risk	Low risk
Vente et al [36]	High risk	High risk	Low risk	Low risk	High risk	Low risk	Low risk
Derradji et al [32]	High risk	High risk	Low risk	Low risk	Low risk	Low risk	Low risk
Fineberg et al [33]	High risk	High risk	Low risk	Low risk	Low risk	Low risk	Low risk

AI in GA Assessment and Progression

Twenty studies explored AI for GA assessment and progression using noninvasive imaging, published between 2019 and 2025 (Table S2 in Multimedia Appendix 2). As shown in Table 3, these studies covered 11 segmentation [5,11,

37-45], 2 algorithm optimization [46,47], 3 AMD progression classification [48-50], and 3 combined tasks such as identification, segmentation, and quantification [51-53]. One study focused solely on GA quantification [54]. Retrospective analyses accounted for 9 studies [5,40,41,43,44,47,48,50,51], while 7 were model development [11,37-39,45,49,53], and

the remainder were prospective [52,54], comparative [46], or cross-sectional [42]. Geographically, contributions came from China (6/20), the United States (7/20), the United Kingdom (2/20), Australia (2/20), France (1/20), Israel (1/20), and Austria (1/20).

Table 3. Characteristics of studies evaluating artificial intelligence (AI) models for geographic atrophy (GA) assessment and progression using noninvasive retinal imaging.

Author	Study design	Region	Purpose of the study	Source of datasets	Number of patients	Number of images or scans	Model evaluation method	Image modality (Image resolution)	AI algorithms	Outcomes	Performance of models
Pramil et al [5]	Retrospective review of images	United States (Boston)	Segmentation (GA lesions)	The “SWAG-GER” cohort of the non-Exudative Age-Related Macular Degeneration (from New England Eye Center at Tufts Medical Center)	90	126	5-fold cross-validation	SS-OCT ^a (500*500 pixels)	CNN ^b ; U-Net	SEN ^c , SPE ^d , and DICE	• SEN=0.95; SPE=0.91; DSC (vs G1): mean 0.92 (SD 0.11); DSC ^e (vs G2): mean 0.91 (SD 0.12).
Siraz et al [50]	Retrospective comparative study	United States (North Carolina)	Classification (central and noncentral GA)	Atrium Health Wake Forest Baptist	104	355	NR ^f	SD-OCT ^g (224*224 pixels)	CNNs: ResNet50, MobileNetV2, and ViT-B/16	AUROC ^h , F ₁ , and ACC ⁱ	• ResNet50: AUROC: mean 0.545 (SD 0.004), F ₁ : mean 0.431 (SD 0.00); ACC: mean 0.756 (SD 0.00). • MobileNet V2: AUROC: mean 0.521 (SD 0.016), F ₁ : mean 0.432 (SD 0.002); ACC: mean 0.756 (SD 0.00). • ViT-B/16: AUROC: mean 0.718 (SD 0.002), F ₁ : mean 0.602 (SD 0.002)

Author	Study design	Region	Purpose of the study	Source of datasets	Number of patients	Number of images or scans	Model evaluation method	Image modality (Image resolution)	AI algorithms	Outcomes	Performance of models
Arslan et al [43]	Retrospective cohort clinical study	Australia (Victoria)	Segmentation (dry lesion area)	The Center for Eye Research Australia or a private ophthalmology practice diagnosed with GA	51	702	5-fold cross-validation	FAF ^l (768*768 or 1536*1536 pixels)	CNN: U-Net	DSC, DSC loss, SEN, SPE, MAE ^m , ACC, R, and P	0.004); ACC: mean 0.780 (SD 0.005).
											• DSC: mean 0.9780 (SD 0.0124); DSC loss: mean 0.0220 (SD 0.0041); SEN: mean 0.9903 (SD 0.0041); SPE: mean 0.7498 (SD 0.0955); MAE: mean 0.0376 (SD 0.0184); ACC: mean 0.9774 (SD 0.0090); P: mean 9837 (SD 0.0116).
Hu et al [48]	Retrospective clinical study	China (Shenyang)	Classification (dry AMD ^p progression phases)	Shenyang Aier Eye Hospital	338	3401	5-fold cross-validation	SD-OCT (NR)	CNNs: EfficientNetV2, DenseNet169, Xception, and ResNet50NF	ACC, SEN, SPE, F ₁ , Macro-f1, and Kappa	• ACC=97.31 %; SEN=89.25 %; SPE=98.80 %; F ₁ =91.21 %; Macro-f1=92.08 %; Kappa=95.4 5%.

Author	Study design	Region	Purpose of the study	Source of datasets	Number of patients	Number of images or scans	Model evaluation method	Image modality (Image resolution)	AI algorithms	Outcomes	Performance of models
Spaide et al [41]	Retrospective analysis and model comparison	United States (Washington)	Segmentation (GA lesion area)	The SWAGGER cohort from the New England Eye Center at Tufts Medical Center	87	126 scans	5-fold cross-validation	SS-OCT (NR)	CNN: U-Net	DSC	<ul style="list-style-type: none">• U-Net-1: 0.82 (95% CI
											0.78-0.86).
											<ul style="list-style-type: none">• U-Net-Avg: 0.88 (95% CI
Vogl et al [47]	Retrospective analysis	Austria (Vienna)	Identification (GA progression after pegcetacoplan treatment)	The FILLY trial	156	NR	NR	SD-OCT (512*512 pixels)	CNN: U-Net	LPR ^q	<ul style="list-style-type: none">• U-Net-Drop: 0.85-0.91),
											<ul style="list-style-type: none">• U-Net-Drop: 0.90 (95% CI
											0.87-0.93),
Szeskin et al [51]	Retrospective analysis	Israel (Jerusalem)	Identification, quantification (GA lesion)	Datasets D1 and D2 from the Hadassah University Medical Center	D1: 18; D2: 16	NR	4-fold cross-validation	SD-OCT (496*1024 pixels and 496*1536 pixels)	CNN: the custom column classification CNN	AUROC, <i>P</i> , <i>R</i> ⁿ , and <i>F</i> ₁	<ul style="list-style-type: none">• Compared with sham treatment, monthly: -28% (-42.8 to -9.4).
											<ul style="list-style-type: none">• Every other month: -23.9% (-40.2 to -3.0).
											<ul style="list-style-type: none">• AUROC=0.970; (Segment) <i>P</i>: mean 0.84 (SD 0.11); <i>R</i>: mean

Author	Study design	Region	Purpose of the study	Source of datasets	Number of patients	Number of images or scans	Model evaluation method	Image modality (Image resolution)	AI algorithms	Outcomes	Performance of models
Spaide et al [40]	Retrospective analysis	United States (California)	Segmentation (GA lesion area)	Proxima A and B	Proxima A: 154; Proxima B: 183	Proxima A: 497; Proxima B: 940	NR	FAF, NIR [†] (768 *768 pixels)	Multimodal DL [‡] ; U-Net; YNet	DSC and r ²	0.94 (SD 0.03);
											(Lesion) P: mean 0.72 (SD 0.03);
											R: mean 0.91 (SD 0.18);
											• (G1- Ynet)DSC: mean 0.92 (SD 0.09);
											• (G1- Unet)DSC: mean 0.90 (SD 0.09);
											• (G2- Ynet)DSC: mean 0.91 (SD 0.09);
											• (G2- Unet)DSC: mean 0.90 (SD 0.09);
											• (G2- Ynet)DSC: mean 0.91 (SD 0.09);
											• (G2- Unet)DSC: mean 0.90 (SD 0.09);
											• (G2- Ynet)DSC: mean 0.91 (SD 0.09);
Al-kharsan et al [44]	Retrospective analysis	United States (Texas)	Segmentation (GA)	The Retina Consultants of Texas and Retina Vitreous Associates	33; 326	367; 348	5-fold cross-validation	SD-OCT (512*496pixels; 200*1024pixels)	CNN: 3D-to-2D U-Net	DSC and r2	0.959.
											• For Spectralis data, DSC=0.826; r ² =0.906.
											• For Cirrus data, DSC=0.824; r ² =0.883.
Chu et al [52]	Prospective study	United States (Washington)	Identification, segmentation, and quantification (GA)	The University of Miami	70; 20; 25	NR	NR	SS-OCT (512*512 pixels)	CNN: U-Net	DSC, SEN, and SPE	• DSC: mean 0.940 (SD 0.032).
											SEN=100% ; SPE:

Author	Study design	Region	Purpose of the study	Source of datasets	Number of patients	Number of images or scans	Model evaluation method	Image modality (Image resolution)	AI algorithms	Outcomes	Performance of models
Merle et al [54]	Prospective observational study	Australia (Victoria)	Quantification (GA)	The Center for Eye Research Australia	50	NR	NR	SD-OCT; FAF (NR)	CNN: U-Net	Spearman correlation coefficient and SEN	<ul style="list-style-type: none">(OCT - automaticall y)Spearman correlation coefficient= 0.85 (95% CI 0.71-0.91); SEN=0.59.
Yang et al [49]	Model development	China (Shenyang)	Classification (stage of dry AMD progression)	Shenyang Aier Excellence Eye Hospital	1310	16,384	3-fold cross-validation	SD-OCT (NR)	CNNs: ResNet50, EfficientNetB4, MobileNetV3, Xception	ACC, SEN, SPE, and F_1	<ul style="list-style-type: none">ACC(GA): ResNet50=92.35%; EfficientNet B4=93.85%; Xception
Ji et al [37]	Model development	China (Nanjing)	Segmentation (GA lesion area)	Dataset1 and dataset2	8; 54	NR	NR	SD-OCT (224*224 pixels)	Weakly supervised multitask	DSC, IoU ^V , AAD ^w , and CCR ^x	<ul style="list-style-type: none">MobileNet V3=89.64%; Xception=91.16%.ACC (nascent GA): ResNet50=91.56%; EfficientNet B4=89.66%; MobileNet V3=89.43%; Xception=85.22%.DSC: mean 0.862 (SD 0.080); IoU: 0.862 (SD 0.080);

Author	Study design	Region	Purpose of the study	Source of datasets	Number of patients	Number of images or scans	Model evaluation method	Image modality (Image resolution)	AI algorithms	Outcomes	Performance of models
Ma et al [38]	Model development	China (Jinan)	Segmentation (GrA lesion area)	Dataset1 and dataset2	62	NR	5-fold cross-validation	SD-OCT (224*224 pixels)	Weakly supervised model: VGG16	DSC, OR ^y , AAD, CC, and AUROC	mean 0.765 (SD 0.119); AAD: mean 0.090 (SD 0.090); CC: 0.992; DSC: mean 0.847 (SD 0.087); OR: mean 0.744 (SD 0.126); AAD: mean 0.150 (SD 0.149); CC: 0.969; AUROC: 0.933.
Royer et al [39]	Model development	France (Issy-Les-Moulineaux)	Segmentation (GrA lesion area)	the Clinical Imaging Center of the Quinze-Vingts Hospital	18	328	8 different random combinations of 12 series to train the model and 6 for the tests	NIR (256*256 pixels)	Unsupervised neural networks: W-net	F1, P, and R	F1: mean 0.87 (SD 0.07); P: mean 0.90 (SD 0.07); R: mean 0.85 (SD 0.11).
Xu et al [11]	Model development	China (Jinan)	Segmentation (GrA lesion area)	dataset1 and dataset2	8 (test I); 56 (test II)	55 (dataset1); 56 (dataset2)	NR	SD-OCT (1024*512*128pixels; 1024*200*200pixels)	Self-learning algorithm	OR, AAD, and CC	OR: mean 84.48% (SD 11.98%); AAD: mean 11.09% (SD 13.61%); CC: 0.9948.
Zhang et al [53]	Model development	United Kingdom (London)	Segmentation and quantification (GrA)	The FILLY study	200	984	NR	SD-OCT (NR)	CNN: U-Net	DSC, ICC ^z , ACC, SEN, SPE, and F1	Approach 1: ACC=0.91 (95% CI 0.89-0.93); F1=0.94

Author	Study design	Region	Purpose of the study	Source of datasets	Number of patients	Number of images or scans	Model evaluation method	Image modality (Image resolution)	AI algorithms	Outcomes	Performance of models
Liu et al [45]	Model development	China (Wuhan)	Segmentation (GrA)	Wuhan Aier Eye Hospital; the public dataset OCTA500	300	2923	5-fold cross-validation	SD-OCT (512*512 pixels)	Self-learning algorithm (dual-branch image projection network)	Jaccard index, DSC, ACC, P^o , and R	(95% CI
											0.92-0.96);
											SEN=0.99
											(95% CI
											0.97-1.00);
											SPE=0.54
											(95% CI
											0.47-0.61);
											DSC: mean
											0.92 (SD
											0.14);
											ICC=0.94.
											• Approach 2:
											ACC=0.94
											(95% CI
											0.92-0.96);
											F_1 =0.96
											(95% CI
											0.94-0.98);
											SEN=0.98
											(95% CI
											0.96-1.00);
											SPE=0.76
											(95% CI
											0.70-0.82);
											DSC: mean
											0.89 (SD
											0.18); ICC:
											0.91.
											• DSC: mean
											7.03 (SD
											2.73);
											Jaccard
											index: mean
											80.96 (SD
											4.29); ACC:
											mean 91.84

Author	Study design	Region	Purpose of the study	Source of datasets	Number of patients	Number of images or scans	Model evaluation method	Image modality (Image resolution)	AI algorithms	Outcomes	Performance of models
Williamson et al [42]	Cross-sectional study	United Kingdom (London)	Segmentation (GA lesion area)	INSIGHT Health Data Research Hub at Moorfields Eye Hospital	9875 (OCT); 81 (FAF)	NR	NR	3D-OCT; FAF (512*512 pixels)	Self-learning algorithm	PPV ^{aa}	(SD 2.13); P: mean 87.12 (SD 2.34); R: mean 86.56 (SD 2.92); • 0.86 (95% CI 0.79-0.92),
											• 0.86 (95% CI 0.79-0.92),
Safai et al [46]	Comparative analysis	United States (Wisconsin)	Identification (the best AI framework for segmentation of GA)	AREDS ^{2ab} study; the GlaxoSmithKline (GSK) study	271 (AREDS2) ; 100 (GSK)	601 (AREDS2); 156 (GSK)	5-fold cross-validation	FAF (512*512 pixels)	CNNs: UNet, FPN ^{ac} , PSPNet, EfficientNet, ResNet, VGG ^{ad} , mViT ^{ae}	CC and DSC	• FPN_EfficientNet: CC=0.98, DSC=0.931. • FPN_CCes Net: CC=0.98, DSC=0.902. • FPN_VGG: CC=0.98, DSC=0.934. • FPN_mViT: CC=0.99, DSC=0.939. • UNet_EfficientNet: CC=0.98, DSC=0.924. • UNet_CCes Net: CC=0.97, DSC=0.930. • UNet_VGG : CC=0.97, DSC=0.896; UNet_mViT: CC=0.99, DSC=0.938.

Author	Study design	Region	Purpose of the study	Source of datasets	Number of patients	Number of images or scans	Model evaluation method	Image modality (Image resolution)	AI algorithms	Outcomes	Performance of models
											<ul style="list-style-type: none">• PSPNet_EfficientNet: CC=0.93, DSC=0.890.• PSPNet_CCesNet: CC=0.87, DSC=0.877.• PSPNet_VGG: CC=0.95, DSC=0.900.• PSPNet_mViT:

Author	Study design	Region	Purpose of the study	Source of datasets	Number of patients	Number of images or scans	Model evaluation method	Image modality (Image resolution)	AI algorithms	Outcomes	Performance of models
aSS-OCT: swept-source OCT. bCNN: convolutional neural network. cSEN: sensitivity. dSPE: specificity. eDSC: dice similarity coefficient. fNR: not reported. gSD-OCT: spectral domain OCT. hAUROC: area under the receiver operating characteristic curve. iACC: accuracy. jCGA: central geographic atrophy. kNCGA: noncentral geographic atrophy. lFAF: fundus autofluorescence. mMAE: mean absolute error. nR: recall. oP: precision. pAMD: age-related macular degeneration. qLPR: local progression rate. rNIR: near-infrared reflectance. sDL: deep learning. t ² : Pearson correlation coefficient. uOCT: optical coherence tomography. vIoU: intersection over union. wAAD: absolute area difference. xCC: correlation coefficient. yOR: overlap ratio. zICC: intraclass coefficient. aappv: positive predictive value. abAREDS2: Age-Related Eye Disease Study 2. acFPN: Feature Pyramid Network. advGG: Visual Geometry Group. aemViT: Mix Vision Transformer.											
											CC=0.98, DSC=0.889.

Dataset configurations varied: 9 out of 20 studies used training, validation, and test sets [40,41,43-46,50-52]; 11 studies used training and test sets [5,11,37-39,49]; 2 studies used training and validation sets [42,48]; 1 study comprised training, tuning, and internal validation sets [53]; and 2 studies did not specify [47,54]. Across studies, at least 14,064 participants provided image data for analysis. Less than half of the studies (9/20, 45%) provided demographic information, with the average age of participants ranging from 55 to 94 years. Six studies were registered with ClinicalTrials.gov (NCT01342926, NCT02503332, NCT02479386, NCT02399072, and NCT04469140 [5,40,41,46,47,53]). To assess the generalization ability of the DL model, cross-validation methods included 5-fold (8/20 studies [5,38,41,43-46,48]), 4-fold (1/20 study [51]), 3-fold (1/20 study [49]), and other approaches (1/20 study [39]). Nine studies did not report validation specifics.

Multiple imaging modalities supported GA assessment: spectral domain optical coherence tomography (SD-OCT) was most common, followed by swept-source OCT (SS-OCT), 3D-OCT, FAF, and NIR. Preprocessing techniques were widely applied to standardize images and improve model performance. Algorithm architectures varied, with U-Net being the most frequently used. Other approaches included custom CNNs, self-learning algorithms, weakly supervised models, and multimodal networks. For example, Hu et al [48] trained the DL models (ResNet-50, Xception, DenseNet169, and EfficientNetV2), evaluating them on a single fold of the validation dataset, with all F_1 -scores exceeding 90%. Yang [49] proposed an ensemble DL architecture that integrated 4 different CNNs, including ResNet50, EfficientNetB4, MobileNetV3, and Xception, to classify dry AMD progression stages. GA lesions on FAF were automatically segmented using multimodal DL networks (U-Net and Y-Net) fed with FAF and NIR images [40]. In contrast to the multimodal algorithms mentioned above (ie, the examples of DL models), Safai [46] investigated 3 distinct

segmentation architectures along with 4 commonly used encoders, resulting in 12 different AI model combinations to determine the optimal AI framework for GA segmentation on FAF images.

From 20 studies, 42 performance sets were collected. Common metrics included correlation coefficient, mean absolute error, Spearman correlation coefficient, intraclass coefficient, overlap ratio, Pearson correlation coefficient (r^2), Kappa, specificity (SPE), sensitivity (SEN), accuracy, positive predictive value (PPV), F_1 -score, P, R, intersection over union, and dice similarity coefficient (DSC). Regarding the segmentation, classification, identification, and quantification of GA in SD-OCT, 12 studies demonstrated performance comparable to that of clinical experts [11,37,38,44,45, 47-51,53,54]. AI was also capable of efficiently detecting, segmenting, and measuring GA in SS-OCT, 3D-OCT, and FAF images, according to 4 studies [5,41,42,52]. AI for GA segmentation in FAF and NIR images, with clinical data showing good segmentation performance [39,40,43].

We performed a comprehensive assessment of the methodological quality of 16 GA assessment and progression studies encompassing 4 domains (Table 4). Only 8 studies detailed the eligibility criteria in the “patient selection” category, while the others had not been published. Three of the studies [47-49] lacked complete datasets, and 3 others [37,39,54] had small datasets or limited volumes of data. In addition, 3 studies [48,49,54] failed to provide information on image formats or resolutions. Two studies [48,49] were ranked as high risk regarding patient selection since the participants included other types of dry AMD (drusen, nascent GA). In terms of applicability, 18 studies were classified as low risk, while 2 were deemed high risk concerning patient selection. Concerning the “Index test,” only 3 algorithms underwent external validation with a different dataset [40,46,53]. All other items were evaluated as low risk.

Table 4. Methodological quality and applicability summary of geographic atrophy (GA) assessment and progression studies using revised Quality Assessment of Diagnostic Accuracy Studies–Artificial Intelligence (QUAUAS-AI).

Study	Risk of bias			Concerns regarding applicability			
	Patient selection	Index test	Reference standard	Flow and timing	Patient selection	Index test	Reference standard
M Hu [48]	High risk	High risk	Low risk	Low risk	High risk	Low risk	Low risk
JK Yang [49]	High risk	High risk	Low risk	Low risk	High risk	Low risk	Low risk
A Safai [46]	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk
WD Vogl [47]	High risk	High risk	Low risk	Low risk	Low risk	Low risk	Low risk
A Szeskin [51]	High risk	High risk	Low risk	Low risk	Low risk	Low risk	Low risk
ZD Chu [52]	High risk	High risk	Low risk	Low risk	Low risk	Low risk	Low risk
ZX Ji [37]	High risk	High risk	Low risk	Low risk	Low risk	Low risk	Low risk
X Ma [38]	High risk	High risk	Low risk	Low risk	Low risk	Low risk	Low risk
C Royer [39]	High risk	High risk	Low risk	Low risk	Low risk	Low risk	Low risk
T Spaide [40]	High risk	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk
T Spaide [41]	High risk	High risk	Low risk	Low risk	Low risk	Low risk	Low risk

Study	Risk of bias				Concerns regarding applicability		
DJ Williamson [42]	Low risk	High risk	Low risk	Low risk	Low risk	Low risk	Low risk
RB Xu [11]	High risk	High risk	Low risk	Low risk	Low risk	Low risk	Low risk
J Arslan [43]	Low risk	High risk	Low risk	Low risk	Low risk	Low risk	Low risk
V Pramil [5]	Low risk	High risk	Low risk	Low risk	Low risk	Low risk	Low risk
GY Zhang [53]	High risk	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk
DA Merle [54]	High risk	High risk	Low risk	Low risk	Low risk	Low risk	Low risk
H AI-khersan [44]	Low risk	High risk	Low risk	Low risk	Low risk	Low risk	Low risk
S Siraz [50]	Low risk	High risk	Low risk	Low risk	Low risk	Low risk	Low risk
XM Liu [45]	High risk	High risk	Low risk	Low risk	Low risk	Low risk	Low risk

AI in Predicting GA Lesion Area and Progression

Eleven studies used AI for predicting GA lesion growth and progression using noninvasive imaging (Table S3 in [Multimedia Appendix 2](#)). These studies were published between 2021 and 2025, with some information provided in [Table 5](#). The study designs consisted of 6 retrospective studies [55-60], 2 model development studies [61,62], 2 post hoc analyses [63,64], and 1 clinical evaluation of a DL algorithm [65]. Participants or images came from various regions: 6 studies were based in the United States [55,57-60,62], 3 in Australia [63-65], 1 in Switzerland [56], and another involving multiple centers in China and the United States [61]. Research aims focused on GA growth prediction [55,56,59-61,63,65], combined prediction and evaluation of lesion features [57], treatment response assessment [58], and integrated segmentation-prediction tasks [62,64].

Table 5. Characteristics of studies evaluating artificial intelligence (AI) models for geographic atrophy (GA) prediction using noninvasive retinal imaging.

Author	Study design	Region	Purpose of the study	Source of datasets	Number of patients	Number of images or scans or cubes	Model evaluation method	Image modality (resolution)	AI algorithms	Outcomes	Performance of models
Gigon et al [56]	Retrospective monocentric study	Switzerland (Lausanne)	Prediction (RORA ^a progression)	Jules Gonin Eye Hospital	119	NR	NR ^b	SD-OCT ^c (384*384 pixels)	CNN ^d ; EfficientNet-b3	DSC ^e	<ul style="list-style-type: none">• 0-6 months: 0.84• 6-12 months: 0.84• >12 months: 0.89
Dow et al [55]	Retrospective cohort study	United States (Atlanta, Georgia, Portland, Oregon, North Carolina; Maryland, Raleigh, Morrisville, Cary); United Kingdom (Durham, South Durham)	Prediction (iAMD ^f to GA within 1 year)	3 independent datasets from AREDS2 ^g and a tertiary referral center and associated satellites	316; 53; 48	1085; 53; 48	5-fold cross-validation	SD-OCT (512 * 1000 pixels)	CNN: Inception v3	SEN ^h , SPE ⁱ , PPV ^j , NPV ^k , ACC ^l	<ul style="list-style-type: none">• SEN: 0.91 (95% CI 0.74-0.98); SPE: 0.80 (95% CI 0.63-0.91); PPV: 0.78 (95% CI 0.70-0.85); NPV: 0.92 (95% CI 0.90-0.95); ACC: 0.85 (95% CI 0.87-0.91)
Cluceru et al [57]	Retrospective clinical study; observation study	United States (California)	Prediction and evaluation (GA growth rate and GA features related to shape and size)	The lampalizumab phase 3 clinical trials and an accompanying observational study	1041; 255	NR	5-fold cross-validation	FAF ^m (384 * 384 pixels)	CNN: VGG16	r^{2n}	<ul style="list-style-type: none">• Full FAF images: 0.44 (95% CI 0.36-0.49)• Rim only: 0.37 (95% CI 0.35-0.4)• Lesion only: 0.34 (95% CI 0.31-0.36)• Background only: 0.3

Author	Study design	Region	Purpose of the study	Source of datasets	Number of patients	Number of images or scans or cubes	Model evaluation method	Image modality (resolution)	AI algorithms	Outcomes	Performance of models
Anecondi et al [58]	Retrospective clinical study; observation study	United States (California)	Prediction and prognosis (GA lesion area and GA growth rate after lampalizumab treatment)	The lampalizumab phase 3 clinical trials and an accompanying observational study	1279; 443; 106; 169	NR	5-fold cross-validation	SD-OCT, FAF (512*512 pixels)	CNN; Inception v3	r^2	(95% CI 0.27-0.33) <ul style="list-style-type: none">Mask only: 0.27 (95% CI 0.24-0.29)
											GA prediction: <ul style="list-style-type: none">FAF-only: 0.98 (95% CI 0.97-0.99)OCT-only: 0.91 (95% CI 0.87-0.95),Multimodal: 0.94 (95% CI 0.92-0.96), GA growth rate: <ul style="list-style-type: none">FAF-only: 0.65 (95% CI 0.52-0.75),OCT-only: 0.36 (95% CI 0.29-0.43),Multimodal: 0.47 (95% CI 0.40-0.54)
Salvi et al [59]	Retrospective analysis	United States (California)	Prediction (the 1 year region of growth of GA lesions)	The following lampalizumab clinical trials and	597	NR	NR	FAF (768*768 pixels or	CNN; U-Net	P ^o , R ^p , DSC, r ²	Whole lesion: <ul style="list-style-type: none">P: mean 0.70 (SD 0.12); R:

Author	Study design	Region	Purpose of the study	Source of datasets	Number of patients	Number of images or scans or cubes	Model evaluation method	Image modality (resolution)	AI algorithms	Outcomes	Performance of models
Yoshida [60]	Retrospective analysis	United States (California)	Prediction (GA progression)	Three prospective clinical trials	1219; 442	NR	5-fold cross-validation	3D OCT (496*1024*49 voxels)	CNNs: (1) en-face intensity maps; (2) SLIVER-net; (3) a 3D CNN; and (4) en-face layer thickness and between-layer intensity maps from a segmentation model	r^2	mean 0.73 (SD 0.12); DSC: mean 0.70 (SD 0.09); r^2 : 0.79
											<ul style="list-style-type: none">GA lesion area: En-face intensity map: 0.91; SLIVER-net: 0.83; 3D DenseNet: 0.90; OCT EZ^q and RPE^r thickness map: 0.90; GA growth rate: En-face intensity map: 0.33; SLIVER-net: 0.33; 3D DenseNet: 0.35; OCT EZ and RPE thickness map: 0.35.ACC: 0.48; Kappa: 0.23; concordance index: 0.69
GS Reiter [63]	Post hoc analysis	Austria (Vienna)	Prediction (GA lesions progression)	the phase II randomized controlled trial FILLY	134	268 scans	5-fold cross-validation	FAF, NIR ^s , SD-OCT (NR)	CNN: PSC-UNet	ACC, Kappa, concordance index	

Author	Study design	Region	Purpose of the study	Source of datasets	Number of patients	Number of images or scans or cubes	Model evaluation method	Image modality (resolution)	AI algorithms	Outcomes	Performance of models
J Mai [64]	Post hoc analysis	Austria (Vienna)	Segmentation, quantification, and prediction (GA lesion and progression)	The phase 2 FILLY clinical trial and the Medical University of Vienna (MUV)	113; 100	226; 967	5-fold cross-validation	SD-OCT, FAF (768*768 and 1536*1536 pixels)	CNN; U-Net	DSC, Hausdorff distance, ICC ^t	<ul style="list-style-type: none">MUV; DSC: mean 0.86 (SD 0.12); Hausdorff distance: mean 0.54 (SD 0.45); FILLY: DSC: mean 0.91 (SD 0.05); Hausdorff distance: mean 0.38 (SD 0.40)
YH Zhang [61]	Model development	China (Nanjing); United States (California)	Prediction (GA growth)	The Byers Eye Institute of Stanford University; the Jiangsu Provincial People's Hospital	22; 3	86 cubes; 33 cubes	Leave-one-out cross-validation	SD-OCT (178*270 pixels)	Recurrent neural network; the bi-directional long-short term memory network; CNN; 3D-UNet	DSC, CC ^u	<ul style="list-style-type: none">Scenario I: DSC: 0.86; CC: 0.83;Scenario II: DSC: 0.89; CC: 0.84;Scenario III: DSC: 0.89; CC: 0.86;Scenario IV: DSC: 0.92; CC: 0.88;Scenario V: DSC: 0.88; CC: 0.85;Scenario VI: DSC: 0.90; CC: 0.86
SX Wang [62]	Model development	United States (California)	Segmentation and prediction (GA lesion area)	The University of California—Los Angeles	147	NR	8-fold cross-validation	SD-OCT, FAF	CNN; U-Net	SEN, SPE, ACC, OR ^v	<ul style="list-style-type: none">ACC: 0.95; SEN: 0.60;

Author	Study design	Region	Purpose of the study	Source of datasets	Number of patients	Number of images or scans or cubes	Model evaluation method	Image modality (resolution)	AI algorithms	Outcomes	Performance of models
J Mai [65]	Clinical evaluation of a DL-based algorithm	Austria (Vienna)	Prediction (GA and GA progression) Lesions progression	The Medical University of Vienna	100	967	5-fold cross-validation	(512*512 pixels) SD-OCT, FAF (NR)	CNN: PSC-UNet	DSC, MAE ^w , and r2	<div><div>SPE: 0.96; OR: 0.65</div><div><div>• 0-1 year: DSC: mean 0.25 (SD 0.16); MAE: mean 0.13 (SD 0.11)</div><div><div>• 1-2 years: DSC: mean 0.38 (SD 0.20); MAE: mean 0.25 (SD 0.24);</div><div><div>• 2-3 years: DSC: mean 0.38 (SD 0.21); MAE: mean 0.35 (SD 0.34);</div><div><div>• >3 years: DSC: mean 0.37 (SD 0.23); MAE:</div></div></div></div></div></div>

Author	Study design	Region	Purpose of the study	Source of datasets	Number of patients	Number of images or scans or cubes	Model evaluation method	Image modality (resolution)	AI algorithms	Outcomes	Performance of models
mean 0.72 (SD 0.48)											
^a RORA: retinal pigment epithelial and outer retinal atrophy. ^b NR: not reported. ^c OCT: optical coherence tomography. ^d CNN: convolutional neural network. ^e DSC: dice similarity coefficient. ^f AMD: age-related macular degeneration. ^g AREDS2: Age-Related Eye Disease Study 2. ^h SEN: sensitivity. ⁱ SPE: specificity. ^j PPV: positive predictive value. ^k NPV: negative predictive value. ^l ACC: accuracy. ^m FAF: fundus autofluorescence. ⁿ χ^2 : Pearson correlation coefficient. ^o P: precision. ^p R: recall. ^q EZ: ellipsoid zone. ^r RPE: retinal pigment epithelium. ^s NIR: near-infrared reflectance. ^t ICC: intraclass coefficient. ^u CC: correlation coefficient. ^v OR: overlap ratio. ^w MAE: mean absolute error.											

Dataset structures varied: 3 out of 11 studies used training-validation-test splits [59,62,64]; 2 out of 11 studies used training-test sets [56,61]; 3 out of 11 studies used training-validation sets [55,63,65]; and the rest adopted development-holdout [57,60] or development-holdout-independent test configurations [58]. In total, 6706 participants were included across studies. Fewer than half of the studies (4/11, 36.4%) reported demographic information, with mean age ranges spanning from 74 to 83 years [55,58,64,65]. Six studies [57-60,63,64] were ethically approved and registered on ClinicalTrials.gov under the following identifiers: NCT02503332, NCT02247479, NCT02247531, NCT02479386, NCT01229215, and NCT02399072. The DL model's generalizability was assessed using leave-one-out cross-validation in 1 study [61], 5-fold cross-validation in 7 studies [55,57,58,60,63-65], and 8-fold cross-validation in 1 study [62]. The remaining 2 studies [56,59] did not specify the cross-validation methodology.

Studies used 3D-OCT, SD-OCT, NIR, and FAF images, primarily sourced from Heidelberg, Zeiss, and Bioptigen devices. While most reported image metrics, 2 studies did not specify resolution details [63,65]. Commonly used DL architectures included Inception v3 [55,58], PSC-UNet [63,65], U-Net [59,62,64], EfficientNet-b3 [56], and VGG16 [57]. In addition, some studies introduced novel approaches, such as en-face intensity maps, SLIVER-net, 3D CNN, and a recurrent neural network, for improved GA progression forecasting.

According to various image modalities, datasets, and follow-up durations, we gathered 31 sets of performance data from 11 studies. The performance metrics included the Hausdorff distance, concordance index, overlap, SEN, SPE, accuracy, mean absolute error, Kappa, DSC, P, PPV, R, r^2 , and negative predictive value. The findings for a single image modality (3D-OCT, SD-OCT, or FAF) demonstrated the development of DL algorithms to predict GA growth rate and progression with excellent performance characteristics comparable to trained experts [55-57,59-61]. Multimodal approaches combining FAF, NIR, and SD-OCT further showed feasibility for individualized lesion growth prediction and localization [58,62-65].

In this systematic review, we used the PROBAST tool to rigorously evaluate prediction models across 4 domains, addressing 20 signaling questions for each paper reviewed. Within the "participants" domain, all studies used appropriate data sources; however, only 6 studies [57-60,64,65] clearly outlined their inclusion and exclusion criteria for participants, leaving the others unclear. In terms of "predictors," these were defined and evaluated similarly for all participants, having no connection to outcome data and being available at baseline. All studies evaluated "yes" to the questions on outcome measurement methods, definitions, interference factors, and measurement time intervals. Concerning "analysis," Dow [55] and Zhang [61] applied a small dataset with an insufficient number of participants. While Zhang performed internal validation, the lack of external validation notably limits the model's generalizability, which was constructed with bi-directional long-short term

memory networks and CNN frameworks. Two studies by Salvi [59] and Yoshida [60] lacked independent and external validation. Gigon [56] failed to explicitly mention missing data handling, complex problems, and model overfitting. Conversely, all other items were evaluated as low risk, and the applications of the studies were universally ranked as low risk (Table S1 in [Multimedia Appendix 3](#)).

Discussion

Principal Findings

This systematic review evaluated the performance of AI, particularly DL algorithms, in detecting and managing GA secondary to dry AMD using noninvasive imaging modalities. Our findings demonstrate that AI models exhibit strong capabilities in accurately detecting, segmenting, quantifying, and predicting GA progression from OCT, FAF, CFP, and NIR imaging, achieving diagnostic accuracy comparable to that of human experts. However, this review also identified several methodological challenges, such as limited sample sizes, inconsistent annotation standards, and a general lack of external validation, which may hinder the clinical generalizability and practical application of these models. Despite these limitations, AI-based tools show significant potential for future use by both specialists and nonspecialists in primary and specialty care settings.

AI in Detecting GA With OCT, FAF, NIR, and CFP Images

Ten studies published between 2018 and 2025 were included, involving at least 7132 participants aged 50 to 85 years. Half of the studies were conducted in the United States, while others originated from European countries. SD-OCT was the most frequently used imaging modality (6/10 studies), followed by CFP (2/10 studies), NIR (1/10 studies), and FAF (1/10 studies). Image preprocessing techniques, such as standardization of size, orientation, and intensity, as well as noise reduction, were consistently applied to enhance model stability and training efficiency. However, 3 studies did not report critical image parameters, such as resolution, potentially limiting reproducibility. DL-based algorithms, including CNNs, were the primary methodologies used for GA detection. Cross-validation techniques, such as 5-fold and 10-fold methods, were used in half of the studies to assess model robustness, though 3 studies did not report validation strategies. AI, particularly DL algorithms, holds significant promise for the detection of GA using noninvasive imaging modalities. OCT, CFP, NIR, and FAF each demonstrated robust diagnostic potential, with performance metrics rivaling or exceeding human expertise.

AI for GA Management With OCT, FAF, and NIR Images

A total of 20 studies (14,064 participants) were published between 2019 and 2025, focusing on themes such as GA segmentation, classification, quantification, and progression prediction. The research designs and geographic regions are diverse. The studies included retrospective analysis (9/20),

model development (7/20), and prospective, comparative, or cross-sectional studies (4/20). Significant contributions came from China (6/20) and the United States (7/20), with additional studies from the United Kingdom (2/20), Australia (2/20), France (1/20), Israel (1/20), and Austria (1/20). The studies used a variety of imaging modalities to assess GA, including SD-OCT, FAF, NIR, SS-OCT, and 3D-OCT. DL algorithms demonstrated remarkable performance in GA management tasks. U-Net was the most commonly used architecture. Multimodal approaches combined FAF and NIR images with DL networks to improve segmentation accuracy. Performance metrics, such as DSC, Kappa, SEN, SPE, and accuracy, consistently showed strong diagnostic accuracy, with several studies achieving performance comparable to clinical experts.

Eleven studies with 6706 participants, published between 2021 and 2025, concentrated on the application of AI for predicting and segmenting GA lesions, as well as their growth and progression. The methodologies were diverse, including retrospective studies, model development studies, post hoc analyses, and clinical algorithm assessment. Participants or images were gathered from regions such as the United States, Australia, Switzerland, and various centers in China and the United States, ensuring broad geographic representation. Demographic information was reported in fewer than half of the studies, with a mean age ranging from 74 to 83 years. Imaging modalities, such as 3D-OCT, SD-OCT, NIR, and FAF, were obtained from devices including Bioptigen, Heidelberg Spectralis HRA+OCT, and Cirrus OCT. While the image preprocessing parameters were consistent across most studies, some did not specify image resolution. Multiview CNN architectures and advanced frameworks, such as the bi-directional long-short term memory networks, were used. DL algorithms exhibited excellent predictive capabilities, with multimodal approaches enabling individualized GA lesion growth prediction.

Noninvasive Image Analysis Techniques for GA

GA, a late-stage form of dry AMD, is marked by the irreversible loss of photoreceptors, RPE, and choriocapillaris [4,5]. The application of noninvasive imaging modalities has revolutionized the detection and management of GA. A comparative summary of AI performance across these modalities is provided in Table S2 in [Multimedia Appendix 3](#). CFP serves as a standard initial assessment tool, useful for screening and early detection. It identifies GA lesions as visible underlying choroidal vessels and well-defined regions of RPE hypopigmentation [66]. FAF imaging using a blue excitation wavelength (488 nm) visualizes metabolic changes at the level of photoreceptor or RPE complex and is practical in assessing GA lesion size and progression with hypo-autofluorescence [67]. In contrast to nonatrophic areas, GA lesions on NIR (787-820 nm, longer than FAF) typically appear brighter and less harmful to the eye [68]. In addition, NIR can help detect the boundaries of foveal lesions, where image contrast is lower on FAF [68]. Recently, the Classification of Atrophy Meeting group recommended that

atrophy in both patients with and those without neovascular AMD be defined based on specific drusen characteristics and other anatomical features, and it is most easily characterized by OCT [69,70]. OCT stands out as the gold standard for GA detection and classification, providing high-resolution, cross-sectional, and en face images of the retina and choroid. SD-OCT is widely used in research and clinical trials, offering precise measurement of GA area and growth rates, while SS-OCT and 3D-OCT offer superior structural insights and potential for AI-driven automation [41,71,72]. Despite the higher cost and technical complexity of advanced OCT technologies, their detailed GA assessment capabilities make them indispensable tools in both clinical practice and research. Furthermore, OCT provides volumetric (3D) structural data, unlike the 2D en face projections of FAF, CFP, and NIR. It allows AI to learn not just the surface appearance of atrophy but also the cross-sectional structure alterations that define and precede GA [3]. As technology advances, the integration of AI and further developments in imaging techniques are expected to enhance the utility of these modalities, overcoming current limitations and expanding their applications in ophthalmology.

Advantages and Challenges of AI Architectures in Clinical Workflow

AI addresses critical limitations of traditional GA monitoring, such as labor-intensive manual grading and intergrader variability [73]. Therefore, automated algorithms enable rapid, standardized analysis of large fundus image datasets, reducing clinician workload and enhancing reproducibility [74]. Furthermore, our review revealed a clear trend in the choice of model architectures tailored to specific clinical tasks. A critical analysis of these architectures is provided in Table S3 in [Multimedia Appendix 3](#). Interestingly, with the advancement of AI algorithm architectures, numerous studies have emerged that use these technologies to identify atrophy caused by various retinal diseases and to evaluate treatment outcomes through image analysis. Miere et al [75] pretrained a DL-based classifier to automatically distinguish GA from atrophy secondary to inherited retinal diseases on FAF according to etiology, using 2 approaches (a trained and validated method and a 10-fold cross-validation method), achieving good accuracy and excellent area under the receiver operating characteristic (AUROC) values. In addition, a study examined the association between treatment and changes in photoreceptor lamina thickness in patients with GA secondary to AMD. The effect of pegcetacoplan on photoreceptors in OCT was supported by this post hoc analysis, which demonstrated that treatment with the drug was linked with reduced outer retinal thinning [76]. Similarly, DL-based OCT image analysis assessed the therapeutic effectiveness of complement component 3 inhibition in delaying GA progression, with findings indicating decreased photoreceptor thinning and loss [77]. Recent studies demonstrating the application of AI algorithms in imaging further validate their potential as reliable supplements to human expertise in the diagnosis and management of GA.

Technical Challenges and Limitations

Despite the promising advancements in AI for GA detection and management, several technical challenges and limitations persist. A significant limitation of OCT-based AI models is their difficulty in distinguishing GA secondary to AMD from other forms of retinal atrophy; thus, the findings may not generalize to broader AMD cases or other retinal diseases, which limits their clinical applicability. In addition, images from different OCT devices show significant variability and imprecision, not offering good enough data acquisition [74]. Another major challenge is the variability in algorithm performance caused by differences in training data, image acquisition protocols, and disease definitions. These differences reduce reproducibility and limit practical deployment. For instance, the absence of standardized reporting in AI studies can result in discrepancies when interpreting results and hinder comparisons between different models. Moreover, despite the high-performance metrics (eg, SEN, SPE, DSC > 0.85, and AUROC > 0.95) reported by many studies, methodological limitations remain. All diagnostic studies included in this review were assessed as high risk in at least 1 domain (10/10), only 1 GA assessment study (1/20) was evaluated as low risk across all domains, and several prediction studies (7/11) were ranked as high or unclear risk in at least 1 domain, primarily due to small or nonrepresentative datasets and a lack of detailed reporting on image preprocessing and external validation. These methodological shortcomings may lead to an overestimation of AI model performance and reduced overall robustness, thereby decreasing the generalizability of the findings and limiting confidence in their real-world applicability. Future studies should prioritize the use of larger, more diverse datasets and implement rigorous validation frameworks to enhance performance metrics (including detection, segmentation, quantification, and prediction accuracy) and conduct prospective, multicenter validation studies to improve clinical applicability and generalizability. Furthermore, adherence to established reporting guidelines for AI studies (such as the Standards for Reporting Diagnostic Accuracy-AI and Checklist for Artificial Intelligence in Medical Imaging [78, 79]) would improve comprehension and transparency, allow for more meaningful comparisons between systems, and facilitate meta-analyses.

Real-World Implications and Research Contributions

Overall, despite some limitations, AI is constantly evolving and holds great potential for transformation in the

health care sector [80]. AI has the potential to accelerate existing forms of medical analysis; however, its algorithms require further testing to be fully trusted. Clinically, AI-based automated tools show strong potential to facilitate early detection, precise quantification, progression, and prediction of GA, thereby reducing the burden on retinal specialists and improving diagnostic consistency. Furthermore, DL algorithms have demonstrated effectiveness in identifying retinal image features associated with cognitive decline, dementia, Parkinson disease, and cardiovascular risk factors [81]. These findings indicate that AI-based retinal images hold promise for transforming primary care and systemic disease management. Although most AI applications remain in the validation phase, the integration of AI with multimodal imaging, novel biomarkers, and emerging therapeutics holds promise for transforming clinical management paradigms in GA and advancing personalized medicine. Future efforts should focus on developing standardized datasets, improving algorithmic generalizability, and conducting real-world validation studies to fully integrate AI into routine ophthalmic practice.

Conclusion

AI, especially DL-based algorithms, holds considerable promise for the detection and management of GA secondary to dry AMD, with performance comparable to trained experts. This systematic review synthesizes and critically appraises the current evidence, highlighting that AI's capabilities extend across GA management—from initial detection and precise segmentation to the forecasting of lesion progression, which informs future research directions. Meanwhile, with the development of C5 inhibitors, AI-based noninvasive fundus image analysis is expected to detect, identify, and monitor GA at an early stage, thereby increasing the window of opportunity in the future. AI has strong potential to augment and streamline clinical workflows by offering automated, reproducible analysis that can assist clinicians in managing large volumes of imaging data; however, more studies are needed to further validate its effectiveness, repeatability, and accuracy.

Acknowledgments

The authors declared that artificial intelligence (AI) or AI-assisted technologies were not used in the writing process of this manuscript.

Funding

This research was funded by the Central High-Level Traditional Chinese Medicine Hospital Project of the Eye Hospital, China Academy of Chinese Medical Sciences (grant no GSP5-82); the National Natural Science Foundation of China (grant no 82274589); the Science and Technology Innovation Project, China Academy of Chinese Medical Sciences (grant no

CI2023C008YG); the Institute-level Research Launch Fund of the Eye Hospital, China Academy of Chinese Medical Sciences (grant no kxy-202402); and the Special Project for the Director of the Business Research Office (grant no 2020YJSZX-2).

Data Availability

All data generated or analyzed during this study are included in this published article and its multimedia appendix files.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Search strategies in all databases.

[\[DOCX File \(Microsoft Word File\), 15 KB-Multimedia Appendix 1\]](#)

Multimedia Appendix 2

Summary of studies on geographic atrophy (GA) detection, GA assessment, GA prediction, quality assessment using Quality Assessment of Diagnostic Accuracy Studies–Artificial Intelligence (QUADAS-AI), and risk of bias evaluation using the Prediction Model Risk of Bias Assessment Tool (PROBAST).

[\[XLSX File \(Microsoft Excel File\), 51 KB-Multimedia Appendix 2\]](#)

Multimedia Appendix 3

Risk of bias and applicability assessment for geographic atrophy (GA) prediction studies, comparative performance of artificial intelligence (AI) models and features across imaging modalities for GA detection and management, and comparison of AI model architectures in GA detection and management.

[\[XLSX File \(Microsoft Excel File\), 15 KB-Multimedia Appendix 3\]](#)

Checklist 1

PRISMA 2020 checklist.

[\[DOCX File \(Microsoft Word File\), 29 KB-Checklist 1\]](#)

Checklist 2

PRISMA-S checklist.

[\[DOCX File \(Microsoft Word File\), 17 KB-Checklist 2\]](#)

References

1. Thomas CJ, Mirza RG, Gill MK. Age-related macular degeneration. *Med Clin North Am*. May 2021;105(3):473-491. [doi: [10.1016/j.mcna.2021.01.003](https://doi.org/10.1016/j.mcna.2021.01.003)] [Medline: [33926642](https://pubmed.ncbi.nlm.nih.gov/33926642/)]
2. Fleckenstein M, Schmitz-Valckenberg S, Chakravarthy U. Age-related macular degeneration: a review. *JAMA*. Jan 9, 2024;331(2):147-157. [doi: [10.1001/jama.2023.26074](https://doi.org/10.1001/jama.2023.26074)] [Medline: [38193957](https://pubmed.ncbi.nlm.nih.gov/38193957/)]
3. Sadda SR, Guymer R, Holz FG, et al. Consensus definition for atrophy associated with age-related macular degeneration on OCT: classification of atrophy report 3. *Ophthalmology*. Apr 2018;125(4):537-548. [doi: [10.1016/j.ophtha.2017.09.028](https://doi.org/10.1016/j.ophtha.2017.09.028)] [Medline: [29103793](https://pubmed.ncbi.nlm.nih.gov/29103793/)]
4. Liakopoulos S, von der Emde L, Biller ML, Ach T, Holz FG. Geographic atrophy in age-related macular degeneration. *Dtsch Arztebl Int*. Feb 7, 2025;122(3):82-88. [doi: [10.3238/arztebl.m2025.0003](https://doi.org/10.3238/arztebl.m2025.0003)] [Medline: [39836449](https://pubmed.ncbi.nlm.nih.gov/39836449/)]
5. Pramil V, de Sisternes L, Omlor L, et al. A deep learning model for automated segmentation of geographic atrophy imaged using swept-source OCT. *Ophthalmol Retina*. Feb 2023;7(2):127-141. [doi: [10.1016/j.oret.2022.08.007](https://doi.org/10.1016/j.oret.2022.08.007)] [Medline: [35970318](https://pubmed.ncbi.nlm.nih.gov/35970318/)]
6. Heier JS, Lad EM, Holz FG, et al. Pegcetacoplan for the treatment of geographic atrophy secondary to age-related macular degeneration (OAKS and DERBY): two multicentre, randomised, double-masked, sham-controlled, phase 3 trials. *Lancet*. Oct 21, 2023;402(10411):1434-1448. [doi: [10.1016/S0140-6736\(23\)01520-9](https://doi.org/10.1016/S0140-6736(23)01520-9)] [Medline: [37865470](https://pubmed.ncbi.nlm.nih.gov/37865470/)]
7. Khanani AM, Patel SS, Staurenghi G, et al. Efficacy and safety of avacincaptad pegol in patients with geographic atrophy (GATHER2): 12-month results from a randomised, double-masked, phase 3 trial. *Lancet*. Oct 21, 2023;402(10411):1449-1458. [doi: [10.1016/S0140-6736\(23\)01583-0](https://doi.org/10.1016/S0140-6736(23)01583-0)] [Medline: [37696275](https://pubmed.ncbi.nlm.nih.gov/37696275/)]
8. Jaffe GJ, Westby K, Csaky KG, et al. C5 inhibitor avacincaptad pegol for geographic atrophy due to age-related macular degeneration: a randomized pivotal phase 2/3 trial. *Ophthalmology*. Apr 2021;128(4):576-586. [doi: [10.1016/j.ophtha.2020.08.027](https://doi.org/10.1016/j.ophtha.2020.08.027)] [Medline: [32882310](https://pubmed.ncbi.nlm.nih.gov/32882310/)]
9. Schmidt-Erfurth U, Mai J, Reiter GS, et al. Disease activity and therapeutic response to pegcetacoplan for geographic atrophy identified by deep learning-based analysis of OCT. *Ophthalmology*. Feb 2025;132(2):181-193. [doi: [10.1016/j.ophtha.2024.08.017](https://doi.org/10.1016/j.ophtha.2024.08.017)] [Medline: [39151755](https://pubmed.ncbi.nlm.nih.gov/39151755/)]

10. Reiter GS, Mai J, Riedl S, et al. AI in the clinical management of GA: a novel therapeutic universe requires novel tools. *Prog Retin Eye Res*. Nov 2024;103:101305. [doi: [10.1016/j.preteyeres.2024.101305](https://doi.org/10.1016/j.preteyeres.2024.101305)] [Medline: [39343193](https://pubmed.ncbi.nlm.nih.gov/39343193/)]
11. Xu R, Niu S, Chen Q, Ji Z, Rubin D, Chen Y. Automated geographic atrophy segmentation for SD-OCT images based on two-stage learning model. *Comput Biol Med*. Feb 2019;105:102-111. [doi: [10.1016/j.compbio.2018.12.013](https://doi.org/10.1016/j.compbio.2018.12.013)] [Medline: [30605812](https://pubmed.ncbi.nlm.nih.gov/30605812/)]
12. Kanagasingam Y, Bhuiyan A, Abramoff MD, Smith RT, Goldschmidt L, Wong TY. Progress on retinal image analysis for age related macular degeneration. *Prog Retin Eye Res*. Jan 2014;38:20-42. [doi: [10.1016/j.preteyeres.2013.10.002](https://doi.org/10.1016/j.preteyeres.2013.10.002)] [Medline: [24211245](https://pubmed.ncbi.nlm.nih.gov/24211245/)]
13. Shi X, Keenan TDL, Chen Q, et al. Improving interpretability in machine diagnosis: detection of geographic atrophy in OCT scans. *Ophthalmol Sci*. Sep 2021;1(3):100038. [doi: [10.1016/j.xops.2021.100038](https://doi.org/10.1016/j.xops.2021.100038)] [Medline: [36247813](https://pubmed.ncbi.nlm.nih.gov/36247813/)]
14. What is artificial intelligence? definition, uses, and types. Coursera. URL: <https://www.coursera.org/articles/what-is-artificial-intelligence> [Accessed 2025-11-07]
15. What is machine learning? definition, types, and examples. Coursera. URL: <https://www.coursera.org/articles/what-is-machine-learning> [Accessed 2025-11-07]
16. What is deep learning? definition, examples, and careers. Coursera. URL: <https://www.coursera.org/articles/what-is-deep-learning> [Accessed 2025-11-07]
17. Aggarwal R, Sounderajah V, Martin G, et al. Diagnostic accuracy of deep learning in medical imaging: a systematic review and meta-analysis. *NPJ Digit Med*. Apr 7, 2021;4(1):65. [doi: [10.1038/s41746-021-00438-z](https://doi.org/10.1038/s41746-021-00438-z)] [Medline: [33828217](https://pubmed.ncbi.nlm.nih.gov/33828217/)]
18. Son J, Shin JY, Kim HD, Jung KH, Park KH, Park SJ. Development and validation of deep learning models for screening multiple abnormal findings in retinal fundus images. *Ophthalmology*. Jan 2020;127(1):85-94. [doi: [10.1016/j.ophtha.2019.05.029](https://doi.org/10.1016/j.ophtha.2019.05.029)] [Medline: [31281057](https://pubmed.ncbi.nlm.nih.gov/31281057/)]
19. McInnes MDF, Moher D, Thombs BD, et al. Preferred Reporting Items for a Systematic Review and Meta-analysis of Diagnostic Test Accuracy Studies: the PRISMA-DTA statement. *JAMA*. Jan 23, 2018;319(4):388-396. [doi: [10.1001/jama.2017.19163](https://doi.org/10.1001/jama.2017.19163)] [Medline: [29362800](https://pubmed.ncbi.nlm.nih.gov/29362800/)]
20. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*. Mar 29, 2021;372:n71. [doi: [10.1136/bmj.n71](https://doi.org/10.1136/bmj.n71)] [Medline: [33782057](https://pubmed.ncbi.nlm.nih.gov/33782057/)]
21. Kang C, Lo JE, Zhang H, et al. Artificial intelligence for diagnosing exudative age-related macular degeneration. *Cochrane Database Syst Rev*. Oct 17, 2024;10(10):CD015522. [doi: [10.1002/14651858.CD015522.pub2](https://doi.org/10.1002/14651858.CD015522.pub2)] [Medline: [39417312](https://pubmed.ncbi.nlm.nih.gov/39417312/)]
22. Rethlefsen ML, Kirtley S, Waffenschmidt S, et al. PRISMA-S: an extension to the PRISMA statement for Reporting Literature Searches in Systematic Reviews. *Syst Rev*. Jan 26, 2021;10(1):39. [doi: [10.1186/s13643-020-01542-z](https://doi.org/10.1186/s13643-020-01542-z)] [Medline: [33499930](https://pubmed.ncbi.nlm.nih.gov/33499930/)]
23. Shi NN, Li J, Liu GH, Cao MF. Artificial intelligence for the detection of glaucoma with SD-OCT images: a systematic review and Meta-analysis. *Int J Ophthalmol*. 2024;17(3):408-419. [doi: [10.18240/ijo.2024.03.02](https://doi.org/10.18240/ijo.2024.03.02)] [Medline: [38721504](https://pubmed.ncbi.nlm.nih.gov/38721504/)]
24. Sounderajah V, Ashrafian H, Rose S, et al. A quality assessment tool for artificial intelligence-centered diagnostic test accuracy studies: QUADAS-AI. *Nat Med*. Oct 2021;27(10):1663-1665. [doi: [10.1038/s41591-021-01517-0](https://doi.org/10.1038/s41591-021-01517-0)]
25. Mohammad-Rahimi H, Motamedian SR, Rohban MH, et al. Deep learning for caries detection: a systematic review. *J Dent (Shiraz)*. Jul 2022;122:104115. [doi: [10.1016/j.jdent.2022.104115](https://doi.org/10.1016/j.jdent.2022.104115)]
26. Moons KGM, Wolff RF, Riley RD, et al. PROBAST: a tool to assess risk of bias and applicability of prediction model studies: explanation and elaboration. *Ann Intern Med*. Jan 1, 2019;170(1):W1-W33. [doi: [10.7326/M18-1377](https://doi.org/10.7326/M18-1377)] [Medline: [30596876](https://pubmed.ncbi.nlm.nih.gov/30596876/)]
27. Chiang JN, Corradetti G, Nittala MG, et al. Automated identification of incomplete and complete retinal epithelial pigment and outer retinal atrophy using machine learning. *Ophthalmol Retina*. Feb 2023;7(2):118-126. [doi: [10.1016/j.oret.2022.08.016](https://doi.org/10.1016/j.oret.2022.08.016)] [Medline: [35995411](https://pubmed.ncbi.nlm.nih.gov/35995411/)]
28. Elsayy A, Keenan TDL, Chen Q, et al. Deep-GA-Net for accurate and explainable detection of geographic atrophy on OCT scans. *Ophthalmol Sci*. Dec 2023;3(4):100311. [doi: [10.1016/j.xops.2023.100311](https://doi.org/10.1016/j.xops.2023.100311)] [Medline: [37304045](https://pubmed.ncbi.nlm.nih.gov/37304045/)]
29. Yao H, Wu Z, Gao SS, et al. Deep learning approaches for detecting of nascent geographic atrophy in age-related macular degeneration. *Ophthalmol Sci*. 2024;4(3):100428. [doi: [10.1016/j.xops.2023.100428](https://doi.org/10.1016/j.xops.2023.100428)] [Medline: [38284101](https://pubmed.ncbi.nlm.nih.gov/38284101/)]
30. Treder M, Lauermann JL, Eter N. Deep learning-based detection and classification of geographic atrophy using a deep convolutional neural network classifier. *Graefes Arch Clin Exp Ophthalmol*. Nov 2018;256(11):2053-2060. [doi: [10.1007/s00417-018-4098-2](https://doi.org/10.1007/s00417-018-4098-2)] [Medline: [30091055](https://pubmed.ncbi.nlm.nih.gov/30091055/)]
31. Kalra G, Cetin H, Whitney J, et al. Machine learning-based automated detection and quantification of geographic atrophy and hypertransmission defects using spectral domain optical coherence tomography. *J Pers Med*. Dec 24, 2022;13(1):37. [doi: [10.3390/jpm13010037](https://doi.org/10.3390/jpm13010037)] [Medline: [36675697](https://pubmed.ncbi.nlm.nih.gov/36675697/)]

32. Derradji Y, Mosinska A, Apostolopoulos S, Ciller C, De Zanet S, Mantel I. Fully-automated atrophy segmentation in dry age-related macular degeneration in optical coherence tomography. *Sci Rep*. Nov 8, 2021;11(1):21893. [doi: [10.1038/s41598-021-01227-0](https://doi.org/10.1038/s41598-021-01227-0)] [Medline: [34751189](https://pubmed.ncbi.nlm.nih.gov/34751189/)]
33. Fineberg A, Tiosano A, Golan N, et al. Near infrared reflectance imaging for the assessment of geographic atrophy using deep learning. *Retina*. Jul 15, 2025. [doi: [10.1097/IAE.0000000000004614](https://doi.org/10.1097/IAE.0000000000004614)] [Medline: [40694826](https://pubmed.ncbi.nlm.nih.gov/40694826/)]
34. Keenan TD, Dharssi S, Peng Y, et al. A deep learning approach for automated detection of geographic atrophy from color fundus photographs. *Ophthalmology*. Nov 2019;126(11):1533-1540. [doi: [10.1016/j.ophtha.2019.06.005](https://doi.org/10.1016/j.ophtha.2019.06.005)] [Medline: [31358385](https://pubmed.ncbi.nlm.nih.gov/31358385/)]
35. Sarao V, Veritti D, De Nardin A, Misciagna M, Foresti G, Lanzetta P. Explainable artificial intelligence model for the detection of geographic atrophy using colour retinal photographs. *BMJ Open Ophthalmol*. Dec 6, 2023;8(1):e001411. [doi: [10.1136/bmjophth-2023-001411](https://doi.org/10.1136/bmjophth-2023-001411)] [Medline: [38057106](https://pubmed.ncbi.nlm.nih.gov/38057106/)]
36. de Vente C, Valmaggia P, Hoyng CB, et al. Generalizable deep learning for the detection of incomplete and complete retinal pigment epithelium and outer retinal atrophy: a MACUSTAR report. *Transl Vis Sci Technol*. Sep 3, 2024;13(9):11. [doi: [10.1167/tvst.13.9.11](https://doi.org/10.1167/tvst.13.9.11)] [Medline: [39235402](https://pubmed.ncbi.nlm.nih.gov/39235402/)]
37. Ji Z, Ma X, Leng T, Rubin DL, Chen Q. Mirrored X-Net: Joint classification and contrastive learning for weakly supervised GA segmentation in SD-OCT. *Pattern Recognit DAGM*. Sep 2024;153:110507. [doi: [10.1016/j.patcog.2024.110507](https://doi.org/10.1016/j.patcog.2024.110507)]
38. Ma X, Ji Z, Niu S, Leng T, Rubin DL, Chen Q. MS-CAM: Multi-Scale Class Activation Maps for weakly-supervised segmentation of geographic atrophy lesions in SD-OCT images. *IEEE J Biomed Health Inform*. Dec 2020;24(12):3443-3455. [doi: [10.1109/JBHI.2020.2999588](https://doi.org/10.1109/JBHI.2020.2999588)] [Medline: [32750923](https://pubmed.ncbi.nlm.nih.gov/32750923/)]
39. Royer C, Sublime J, Rossant F, Paques M. Unsupervised approaches for the segmentation of dry ARMD lesions in eye fundus cSLO images. *J Imaging*. Aug 11, 2021;7(8):143. [doi: [10.3390/jimaging7080143](https://doi.org/10.3390/jimaging7080143)] [Medline: [34460779](https://pubmed.ncbi.nlm.nih.gov/34460779/)]
40. Spaide T, Jiang J, Patil J, et al. Geographic atrophy segmentation using multimodal deep learning. *Trans Vis Sci Tech*. Jul 3, 2023;12(7):10. [doi: [10.1167/tvst.12.7.10](https://doi.org/10.1167/tvst.12.7.10)]
41. Spaide T, Rajesh AE, Gim N, et al. Estimating uncertainty of geographic atrophy segmentations with Bayesian deep learning. *Ophthalmol Sci*. 2025;5(1):100587. [doi: [10.1016/j.xops.2024.100587](https://doi.org/10.1016/j.xops.2024.100587)] [Medline: [39380882](https://pubmed.ncbi.nlm.nih.gov/39380882/)]
42. Williamson DJ, Struyven RR, Antaki F, et al. Artificial intelligence to facilitate clinical trial recruitment in age-related macular degeneration. *Ophthalmol Sci*. 2024;4(6):100566. [doi: [10.1016/j.xops.2024.100566](https://doi.org/10.1016/j.xops.2024.100566)] [Medline: [39139546](https://pubmed.ncbi.nlm.nih.gov/39139546/)]
43. Arslan J, Samarasinghe G, Sowmya A, et al. Deep learning applied to automated segmentation of geographic atrophy in fundus autofluorescence images. *Trans Vis Sci Tech*. Jul 1, 2021;10(8):2. [doi: [10.1167/tvst.10.8.2](https://doi.org/10.1167/tvst.10.8.2)]
44. Al-kharsan H, Oakley J, Russakoff D, et al. Automated deep learning-based 3D-to-2D segmentation of geographic atrophy in optical coherence tomography data. *medRxiv*. Preprint posted online on 2025. [doi: [10.1101/2025.07.07.25330902](https://doi.org/10.1101/2025.07.07.25330902)]
45. Liu X, Li J, Zhang Y, Yao J. Dual-branch image projection network for geographic atrophy segmentation in retinal OCT images. *Sci Rep*. 2025;15(1). [doi: [10.1038/s41598-025-90709-6](https://doi.org/10.1038/s41598-025-90709-6)]
46. Safai A, Froines C, Slater R, et al. Quantifying geographic atrophy in age-related macular degeneration: a comparative analysis across 12 deep learning models. *Invest Ophthalmol Vis Sci*. Jul 1, 2024;65(8):42. [doi: [10.1167/iov.65.8.42](https://doi.org/10.1167/iov.65.8.42)] [Medline: [39046755](https://pubmed.ncbi.nlm.nih.gov/39046755/)]
47. Vogl WD, Riedl S, Mai J, et al. Predicting topographic disease progression and treatment response of pegcetacoplan in geographic atrophy quantified by deep learning. *Ophthalmol Retina*. Jan 2023;7(1):4-13. [doi: [10.1016/j.oret.2022.08.003](https://doi.org/10.1016/j.oret.2022.08.003)] [Medline: [35948209](https://pubmed.ncbi.nlm.nih.gov/35948209/)]
48. Hu M, Wu B, Lu D, et al. Two-step hierarchical neural network for classification of dry age-related macular degeneration using optical coherence tomography images. *Front Med*. 2023;10:1221453. [doi: [10.3389/fmed.2023.1221453](https://doi.org/10.3389/fmed.2023.1221453)]
49. Yang J, Wu B, Wang J, et al. Dry age-related macular degeneration classification from optical coherence tomography images based on ensemble deep learning architecture. *Front Med*. 2024;11:1438768. [doi: [10.3389/fmed.2024.1438768](https://doi.org/10.3389/fmed.2024.1438768)]
50. Siraz S, Kamanda H, Gholami S, Nabil AS, Yee Ong SS, Alam MN. Multi-class classification of central and non-central geographic atrophy using optical coherence tomography. *medRxiv*. Preprint posted online on May 28, 2025. [doi: [10.1101/2025.05.27.25328446](https://doi.org/10.1101/2025.05.27.25328446)] [Medline: [40492092](https://pubmed.ncbi.nlm.nih.gov/40492092/)]
51. Szeskin A, Yehuda R, Shmueli O, Levy J, Joskowicz L. A column-based deep learning method for the detection and quantification of atrophy associated with AMD in OCT scans. *Med Image Anal*. Aug 2021;72:102130. [doi: [10.1016/j.media.2021.102130](https://doi.org/10.1016/j.media.2021.102130)] [Medline: [34198041](https://pubmed.ncbi.nlm.nih.gov/34198041/)]
52. Chu Z, Wang L, Zhou X, et al. Automatic geographic atrophy segmentation using optical attenuation in OCT scans with deep learning. *Biomed Opt Express*. Mar 1, 2022;13(3):1328-1343. [doi: [10.1364/BOE.449314](https://doi.org/10.1364/BOE.449314)] [Medline: [35414972](https://pubmed.ncbi.nlm.nih.gov/35414972/)]

53. Zhang G, Fu DJ, Liefers B, et al. Clinically relevant deep learning for detection and quantification of geographic atrophy from optical coherence tomography: a model development and external validation study. *Lancet Digit Health*. Oct 2021;3(10):e665-e675. [doi: [10.1016/S2589-7500\(21\)00134-5](https://doi.org/10.1016/S2589-7500(21)00134-5)] [Medline: [34509423](https://pubmed.ncbi.nlm.nih.gov/34509423/)]
54. Merle DA, Guymer RH, Chia MA, et al. Mapping the impact: AI-driven quantification of geographic atrophy on OCT scans and its association with visual sensitivity loss. *Br J Ophthalmol*. Sep 23, 2025;109(10):1187-1193. [doi: [10.1136/bjo-2024-326603](https://doi.org/10.1136/bjo-2024-326603)] [Medline: [40490296](https://pubmed.ncbi.nlm.nih.gov/40490296/)]
55. Dow ER, Jeong HK, Katz EA, et al. A deep-learning algorithm to predict short-term progression to geographic atrophy on spectral-domain optical coherence tomography. *JAMA Ophthalmol*. Nov 1, 2023;141(11):1052-1061. [doi: [10.1001/jamaophthalmol.2023.4659](https://doi.org/10.1001/jamaophthalmol.2023.4659)] [Medline: [37856139](https://pubmed.ncbi.nlm.nih.gov/37856139/)]
56. Gigon A, Mosinska A, Montesel A, et al. Personalized atrophy risk mapping in age-related macular degeneration. *Transl Vis Sci Technol*. Nov 1, 2021;10(13):18. [doi: [10.1167/tvst.10.13.18](https://doi.org/10.1167/tvst.10.13.18)] [Medline: [34767623](https://pubmed.ncbi.nlm.nih.gov/34767623/)]
57. Cluceru J, Anegondi N, Gao SS, et al. Topographic clinical insights from deep learning-based geographic atrophy progression prediction. *Transl Vis Sci Technol*. Aug 1, 2024;13(8):6. [doi: [10.1167/tvst.13.8.6](https://doi.org/10.1167/tvst.13.8.6)] [Medline: [39102242](https://pubmed.ncbi.nlm.nih.gov/39102242/)]
58. Anegondi N, Gao SS, Steffen V, et al. Deep learning to predict geographic atrophy area and growth rate from multimodal imaging. *Ophthalmol Retina*. Mar 2023;7(3):243-252. [doi: [10.1016/j.oret.2022.08.018](https://doi.org/10.1016/j.oret.2022.08.018)] [Medline: [36038116](https://pubmed.ncbi.nlm.nih.gov/36038116/)]
59. Salvi A, Cluceru J, Gao SS, et al. Deep learning to predict the future growth of geographic atrophy from fundus autofluorescence. *Ophthalmol Sci*. 2025;5(2):100635. [doi: [10.1016/j.xops.2024.100635](https://doi.org/10.1016/j.xops.2024.100635)] [Medline: [39758130](https://pubmed.ncbi.nlm.nih.gov/39758130/)]
60. Yoshida K, Anegondi N, Pely A, et al. Deep learning approaches to predict geographic atrophy progression using three-dimensional OCT imaging. *Trans Vis Sci Tech*. Feb 3, 2025;14(2):11. [doi: [10.1167/tvst.14.2.11](https://doi.org/10.1167/tvst.14.2.11)]
61. Zhang Y, Zhang X, Ji Z, et al. An integrated time adaptive geographic atrophy prediction model for SD-OCT images. *Med Image Anal*. Feb 2021;68:101893. [doi: [10.1016/j.media.2020.101893](https://doi.org/10.1016/j.media.2020.101893)] [Medline: [33260118](https://pubmed.ncbi.nlm.nih.gov/33260118/)]
62. Wang S, Wang Z, Vejalla S, et al. Reverse engineering for reconstructing baseline features of dry age-related macular degeneration in optical coherence tomography. *Sci Rep*. 2020;12(1):22620. [doi: [10.1038/s41598-022-27140-8](https://doi.org/10.1038/s41598-022-27140-8)]
63. Reiter GS, Lachinov D, Bühl W, et al. A novel management challenge in age-related macular degeneration: artificial intelligence and expert prediction of geographic atrophy. *Ophthalmol Retina*. May 2025;9(5):421-430. [doi: [10.1016/j.oret.2024.10.029](https://doi.org/10.1016/j.oret.2024.10.029)] [Medline: [39522752](https://pubmed.ncbi.nlm.nih.gov/39522752/)]
64. Mai J, Lachinov D, Riedl S, et al. Clinical validation for automated geographic atrophy monitoring on OCT under complement inhibitory treatment. *Sci Rep*. Apr 29, 2023;13(1):7028. [doi: [10.1038/s41598-023-34139-2](https://doi.org/10.1038/s41598-023-34139-2)] [Medline: [37120456](https://pubmed.ncbi.nlm.nih.gov/37120456/)]
65. Mai J, Lachinov D, Reiter GS, et al. Deep learning-based prediction of individual geographic atrophy progression from a single baseline OCT. *Ophthalmol Sci*. 2024;4(4):100466. [doi: [10.1016/j.xops.2024.100466](https://doi.org/10.1016/j.xops.2024.100466)] [Medline: [38591046](https://pubmed.ncbi.nlm.nih.gov/38591046/)]
66. Fang H, Li F, Fu H, et al. ADAM challenge: detecting age-related macular degeneration from fundus images. *IEEE Trans Med Imaging*. Oct 2022;41(10):2828-2847. [doi: [10.1109/TMI.2022.3172773](https://doi.org/10.1109/TMI.2022.3172773)] [Medline: [35507621](https://pubmed.ncbi.nlm.nih.gov/35507621/)]
67. Mukherjee S, Arunachalam T, Duic C, et al. Structure-function relationships in geographic atrophy based on mesopic microperimetry, fundus autofluorescence, and optical coherence tomography. *Transl Vis Sci Technol*. Feb 3, 2025;14(2):7. [doi: [10.1167/tvst.14.2.7](https://doi.org/10.1167/tvst.14.2.7)] [Medline: [39908134](https://pubmed.ncbi.nlm.nih.gov/39908134/)]
68. Keilhauer CN, Delori FC. Near-infrared autofluorescence imaging of the fundus: visualization of ocular melanin. *Invest Ophthalmol Vis Sci*. Aug 2006;47(8):3556-3564. [doi: [10.1167/iovs.06-0122](https://doi.org/10.1167/iovs.06-0122)] [Medline: [16877429](https://pubmed.ncbi.nlm.nih.gov/16877429/)]
69. Jaffe GJ, Chakravarthy U, Freund KB, et al. Imaging features associated with progression to geographic atrophy in age-related macular degeneration: Classification of Atrophy Meeting Report 5. *Ophthalmol Retina*. Sep 2021;5(9):855-867. [doi: [10.1016/j.oret.2020.12.009](https://doi.org/10.1016/j.oret.2020.12.009)] [Medline: [33348085](https://pubmed.ncbi.nlm.nih.gov/33348085/)]
70. Clevenger L, Rachitskaya A. Identifying geographic atrophy. *Curr Opin Ophthalmol*. May 1, 2023;34(3):195-202. [doi: [10.1097/ICU.0000000000000952](https://doi.org/10.1097/ICU.0000000000000952)] [Medline: [36943458](https://pubmed.ncbi.nlm.nih.gov/36943458/)]
71. Velaga SB, Nittala MG, Hariri A, Sadda SR. Correlation between fundus autofluorescence and en face OCT measurements of geographic atrophy. *Ophthalmol Retina*. Aug 2022;6(8):676-683. [doi: [10.1016/j.oret.2022.03.017](https://doi.org/10.1016/j.oret.2022.03.017)] [Medline: [35338026](https://pubmed.ncbi.nlm.nih.gov/35338026/)]
72. Shi Y, Zhang Q, Zhou H, et al. Correlations between choriocapillaris and choroidal measurements and the growth of geographic atrophy using swept source OCT imaging. *Am J Ophthalmol*. Apr 2021;224:321-331. [doi: [10.1016/j.ajo.2020.12.015](https://doi.org/10.1016/j.ajo.2020.12.015)] [Medline: [33359715](https://pubmed.ncbi.nlm.nih.gov/33359715/)]
73. Wei W, Anantharajit R, Patel RP, Cordeiro MF. Detection of macular atrophy in age-related macular degeneration aided by artificial intelligence. *Expert Rev Mol Diagn*. Jun 2023;23(6):485-494. [doi: [10.1080/14737159.2023.2208751](https://doi.org/10.1080/14737159.2023.2208751)] [Medline: [37144908](https://pubmed.ncbi.nlm.nih.gov/37144908/)]
74. Enzendorfer ML, Schmidt-Erfurth U. Artificial intelligence for geographic atrophy: pearls and pitfalls. *Curr Opin Ophthalmol*. Nov 1, 2024;35(6):455-462. [doi: [10.1097/ICU.0000000000001085](https://doi.org/10.1097/ICU.0000000000001085)] [Medline: [39259599](https://pubmed.ncbi.nlm.nih.gov/39259599/)]

75. Miere A, Capuano V, Kessler A, et al. Deep learning-based classification of retinal atrophy using fundus autofluorescence imaging. *Comput Biol Med.* Mar 2021;130:104198. [doi: [10.1016/j.compbimed.2020.104198](https://doi.org/10.1016/j.compbimed.2020.104198)] [Medline: [33383315](https://pubmed.ncbi.nlm.nih.gov/33383315/)]
76. Pfau M, Schmitz-Valckenberg S, Ribeiro R, et al. Association of complement C3 inhibitor pegcetacoplan with reduced photoreceptor degeneration beyond areas of geographic atrophy. *Sci Rep.* Oct 25, 2022;12(1):17870. [doi: [10.1038/s41598-022-22404-9](https://doi.org/10.1038/s41598-022-22404-9)] [Medline: [36284220](https://pubmed.ncbi.nlm.nih.gov/36284220/)]
77. Riedl S, Vogl WD, Mai J, et al. The effect of pegcetacoplan treatment on photoreceptor maintenance in geographic atrophy monitored by artificial intelligence-based OCT analysis. *Ophthalmol Retina.* Nov 2022;6(11):1009-1018. [doi: [10.1016/j.oret.2022.05.030](https://doi.org/10.1016/j.oret.2022.05.030)] [Medline: [35667569](https://pubmed.ncbi.nlm.nih.gov/35667569/)]
78. Sounderajah V, Guni A, Liu X, et al. The STARD-AI reporting guideline for diagnostic accuracy studies using artificial intelligence. *Nat Med.* Oct 2025;31(10):3283-3289. [doi: [10.1038/s41591-025-03953-8](https://doi.org/10.1038/s41591-025-03953-8)] [Medline: [40954311](https://pubmed.ncbi.nlm.nih.gov/40954311/)]
79. Mongan J, Moy L, Kahn CE. Checklist for Artificial Intelligence in Medical Imaging (CLAIM): a guide for authors and reviewers. *Radiol Artif Intell.* Mar 2020;2(2):e200029. [doi: [10.1148/ryai.2020200029](https://doi.org/10.1148/ryai.2020200029)] [Medline: [33937821](https://pubmed.ncbi.nlm.nih.gov/33937821/)]
80. Wang J, Wang YX, Zeng D, et al. Artificial intelligence-enhanced retinal imaging as a biomarker for systemic diseases. *Theranostics.* 2025;15(8):3223-3233. [doi: [10.7150/thno.100786](https://doi.org/10.7150/thno.100786)]
81. Tan YY, Kang HG, Lee CJ, et al. Correction: prognostic potentials of AI in ophthalmology: systemic disease forecasting via retinal imaging. *Eye and Vis.* 2024;11(1). [doi: [10.1186/s40662-024-00399-w](https://doi.org/10.1186/s40662-024-00399-w)]

Abbreviations

AI: artificial intelligence
AMD: age-related macular degeneration
AUROC: area under the receiver operating characteristic curve
CFP: color fundus photography
CNN: convolutional neural network
cRORA: complete retinal pigment epithelium and outer retinal atrophy
DL: deep learning
DSC: dice similarity coefficient
FAF: fundus autofluorescence
GA: geographic atrophy
ML: machine learning
NIR: near-infrared reflectance
OCT: optical coherence tomography
PPV: positive predictive value
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PRISMA-DTA: Preferred Reporting Items for Systematic Reviews and Meta-Analysis of Diagnostic Test Accuracy
PRISMA-S: Preferred Reporting Items for Systematic Reviews and Meta-Analyses literature search extension
PROBAST: Prediction Model Risk of Bias Assessment Tool
QUADAS: Quality Assessment of Diagnostic Accuracy Studies
RPE: retinal pigment epithelium
SD-OCT: spectral domain optical coherence tomography
SEN: sensitivity
SPE: specificity
SS-OCT: swept-source OCT

Edited by Amaryllis Mavragani, Stefano Brini; peer-reviewed by Jiale Zhang, Xiaolong Liang; submitted 26.Jul.2025; final revised version received 11.Oct.2025; accepted 11.Oct.2025; published 21.Nov.2025

Please cite as:

Shi N, Li J, Shang M, Zhang W, Xu K, Li Y, Liang L

Detection and Management of Geographic Atrophy Secondary to Age-Related Macular Degeneration Using Noninvasive Retinal Images and Artificial Intelligence: Systematic Review

*J Med Internet Res*2025;27:e81328

URL: <https://www.jmir.org/2025/1/e81328>

doi: [10.2196/81328](https://doi.org/10.2196/81328)

© Nannan Shi, Jiaxian Li, Mengqiu Shang, Weidao Zhang, Kai Xu, Yamin Li, Lina Liang. Originally published in the Journal of Medical Internet Research (<https://www.jmir.org>), 21.Nov.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research (ISSN 1438-8871), is properly cited. The complete bibliographic information, a link to the original publication on <https://www.jmir.org/>, as well as this copyright and license information must be included.