



Artificial intelligence in ophthalmology: From diagnostic accuracy to clinical application

Marco Zeppieri, Matteo Capobianco, Federico Visalli, Marieme Khouyyi, Mutali Musa, Alessandro Avitabile, Inferra Leandro, Rosa Giglio, Daniele Tognetto, Caterina Gagliano, Fabiana D'Esposito, Francesco Cappellani

Peer review: Externally peer reviewed.

Peer-review model: Single blind

Peer-review report's classification

Scientific quality: Grade B, Grade C

Novelty: Grade B, Grade C

Creativity or innovation: Grade C, Grade C

Scientific significance: Grade C, Grade D

P-Reviewer: Au SCL, Chief Physician, China

Received: October 15, 2025

Revised: November 12, 2025

Accepted: January 21, 2026

Published online: September 20, 2026

Processing time: 270 Days and 21.3 Hours

Copyright: ©Author(s) 2026. This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution-NonCommercial \(CC BY-NC 4.0\)](#) license. No commercial re-use. See [permissions](#). Published by Baishideng Publishing Group Inc.



Marco Zeppieri, Department of Ophthalmology, University Hospital of Udine, Udine 33100, Italy

Marco Zeppieri, Inferra Leandro, Rosa Giglio, Daniele Tognetto, Department of Medicine, Surgery and Health Sciences, University of Trieste, Trieste 34129, Italy

Matteo Capobianco, Eye Clinic, Azienda Ospedaliero Universitaria Policlinico "G. Rodolico - San Marco" Catania, Catania 95121, Italy

Matteo Capobianco, Alessandro Avitabile, Faculty of Medicine, University of Catania, Catania 95123, Italy

Federico Visalli, Department of Ophthalmology, University of Catania, Catania 95123, Italy

Marieme Khouyyi, Department of Biomedical and Dental Sciences and Morphofunctional Imaging, University of Messina, Messina 98121, Italy

Mutali Musa, Department of Optometry, University of Benin, Benin 300283, Nigeria

Mutali Musa, Department of Ophthalmology, Africa Eye Laser Center Ltd., Benin 300211, Nigeria

Caterina Gagliano, Francesco Cappellani, Department of Medicine and Surgery, Kore University of Enna, Enna 94100, Italy

Caterina Gagliano, Fabiana D'Esposito, Francesco Cappellani, Mediterranean Foundation "G.B. Morgagni", Catania 95125, Italy

Fabiana D'Esposito, Imperial College Ophthalmic Research Group Unit, Imperial College, London NW1 5QH, United Kingdom

ORCID number: Marco Zeppieri [0000-0003-0999-5545](https://orcid.org/0000-0003-0999-5545); Mutali Musa [0000-0001-7486-8361](https://orcid.org/0000-0001-7486-8361); Daniele Tognetto [0000-0001-7197-7765](https://orcid.org/0000-0001-7197-7765); Caterina Gagliano [0000-0001-8424-0068](https://orcid.org/0000-0001-8424-0068); Fabiana D'Esposito [0000-0002-7938-876X](https://orcid.org/0000-0002-7938-876X).

Corresponding author: Caterina Gagliano, MD, PhD, Department of Medicine and Surgery, Kore University of Enna, Viale delle Olimpiadi 1, Enna 94100, Italy.

caterina.gagliano@unikore.it

Abstract

Artificial intelligence in ophthalmology encounters a continual challenge: Systems proficient in picture classification seldom yield quantifiable enhancements in patient outcomes. The primary concern is the disparity between pixel-level performance metrics and their clinical significance. Primary obstacles encompass data bias, domain shift, and label noise, exacerbated by the lack of prospective, randomized deployment trials. The frequent disregard for patient-centered objectives, cost-effectiveness, and equity evaluations is significant. Rectifying these deficiencies necessitates stringent external validation, established decision criteria, and ongoing surveillance within actual clinical practices. Transparent reporting criteria and the deliberate incorporation of human-factors engineering are essential. Only by bridging this gap can algorithmic accuracy be converted into significant diagnostic precision for glaucoma, diabetic retinopathy, and macular conditions (specifically diabetic macular edema and age-related macular degeneration). This paper aims to assess the limits of using high-performing artificial intelligence systems in ocular image processing, which seldom lead to enhanced patient outcomes, and to delineate the scientific, clinical, and practical techniques required to close this gap.

Key Words: Ophthalmology; Artificial intelligence; Image classification; Diagnostic precision; Clinical validation

Core Tip: Artificial intelligence offers potential for enhancing ocular diagnoses; nevertheless, greater algorithmic accuracy at the image level rarely translates to improved patient outcomes. Obstacles, including data bias, domain shift, and label noise, coupled with a scarcity of prospective research and a lack of cost-effectiveness or equity evaluations, impede translation. Enhancing external validation, establishing predetermined thresholds, and incorporating human-factors engineering into healthcare workflows are imperative. Bridging this gap could convert algorithmic accuracy into significant diagnostic precision for glaucoma, diabetic retinopathy, and macular conditions.

Citation: Zeppieri M, Capobianco M, Visalli F, Khouyyi M, Musa M, Avitabile A, Leandro I, Giglio R, Tognetto D, Gagliano C, D'Esposito F, Cappellani F. Artificial intelligence in ophthalmology: From diagnostic accuracy to clinical application. *World J Methodol* 2026; 16(3): 115265

URL: <https://www.wjgnet.com/2222-0682/full/v16/i3/115265.htm>

DOI: <https://dx.doi.org/10.5662/wjm.115265>

INTRODUCTION

Artificial intelligence (AI) has swiftly become a revolutionary influence across medical professions, especially in areas where diagnostic imaging is crucial. Ophthalmology, characterized by its dependence on high-resolution imaging techniques like optical coherence tomography (OCT) and fundus photography, has been a leader in the incorporation of AI in clinical research and practice. The potential of AI lies in its ability to swiftly and accurately analyze extensive and intricate datasets, often surpassing the diagnostic efficacy of experienced doctors in retrospective assessments. Initial pivotal experiments revealed that deep learning algorithms could categorize retinal illnesses, including diabetic retinopathy (DR), glaucoma, and age-related macular degeneration (AMD), with accuracy akin to that of ophthalmologists, thereby paving the way for expanded clinical applications[1].

The allure of AI in ophthalmology transcends mere diagnostic precision. As the global incidence of vision-threatening disorders, notably DR, glaucoma, and AMD, escalates, AI-driven screening and diagnostic technologies are recognized as promising remedies to address workforce deficiencies and mitigate preventable blindness. Regulatory achievements, exemplified by the United States Food and Drug Administration (FDA) endorsement of the IDx-DR autonomous diagnostic system, underscored the preparedness of AI to transition from laboratory environments to clinical applications [2]. Nonetheless, despite these advancements, the tangible effect of AI on patient-centered outcomes has been less persuasive. Although AI models demonstrate proficiency in pixel-level categorization tasks, their capacity to influence patient clinical outcomes through vision preservation, cost reduction, or enhancement of care equity is constrained[3]. A multitude of variables contribute to this translational gap. The predominant AI research in ophthalmology is retrospective and depends on meticulously kept datasets. These datasets, although optimal for algorithm development, frequently exhibit deficiencies in diversity regarding ethnicity, region, disease spectrum, and image quality. Consequently, algorithmic efficacy may diminish when utilized in novel clinical settings[4]. Moreover, deep learning techniques often operate as “black box” systems, producing outputs without clear explanations of their underlying reasoning processes. This opacity erodes physician trust and restricts adoption, especially in contexts where accountability and medicolegal liability are involved[5].

A further area of concern is the absence of prospective, randomized, real-world implementation trials. Despite AI systems demonstrating high accuracy in retrospective analyses, there is a paucity of studies assessing their influence on clinical workflows, cost-effectiveness, or long-term patient outcomes[6]. Furthermore, significant ethical issues, including algorithmic bias, equal access, and data privacy, are often inadequately addressed in the literature[7]. These inadequacies underscore the need for a realignment in AI research priorities, shifting from optimizing algorithmic efficacy on ben-

chmark datasets to demonstrating tangible benefits for patients across diverse healthcare settings.

This invited review aims to achieve two objectives: First, to critically evaluate the limitations of existing AI applications in ophthalmology that impede the conversion of pixel-level diagnostic accuracy into significant patient outcomes; and second, to delineate the scientific, clinical, and implementation strategies necessary to close this gap. This paper seeks to recontextualize the discourse on AI in ophthalmology by emphasizing methodological rigor, clinical validation, and patient-centered metrics, thereby shifting the focus from technical accomplishments to practical precision that directly benefits patients with glaucoma, diabetic eye disease, and macular disorders (specifically diabetic macular edema and AMD).

LITERATURE REVIEW

This study serves as a narrative synthesis of current work regarding AI in ophthalmology, focusing specifically on the disparity between algorithmic precision and patient-centered results. The search technique employed Boolean operators as follows: ('artificial intelligence' OR 'machine learning' OR 'deep learning') AND ('ophthalmology' OR 'retina' OR 'glaucoma' OR 'macular degeneration' OR 'diabetic retinopathy' OR 'OCT' OR 'fundus photography'). Filters were implemented to encompass English-language, human-subject, peer-reviewed studies published from January 2018 to September 2025.

The eligibility criteria for inclusion were extensive, consistent with the narrative character of this study. Studies were included if they presented original data, systematic or narrative reviews, or consensus statements pertaining to AI in ophthalmology. The principal inclusion criteria were: (1) Direct utilization of AI in ocular imaging techniques (fundus, OCT, visual fields, anterior segment imaging); (2) Assessment of diagnostic or prognostic precision; (3) Examination of methodological constraints, including data bias, generalizability, or interpretability; and (4) Clinical or ethical implications pertaining to patient care. The exclusion criteria encompassed solely technical computer science papers lacking clinical significance, animal or *in vitro* investigations, and conference abstracts that did not have peer-reviewed full texts.

The most clinically pertinent papers from the identified literature were picked for comprehensive investigation. These encompassed seminal early investigations that demonstrated the capabilities of AI in ophthalmic diagnosis, essential evaluations of OCT and fundus-based algorithms, systematic reviews regarding AI implementation, and contemporary analyses of bias, trust, and equity in AI integration. The study included studies examining regulatory frameworks, cost-effectiveness calculations, and practical implementation obstacles to ensure rigor. This review, while structured as a narrative synthesis, adheres to the methodological guidelines specified in the Preferred Reporting Items for Systematic reviews and Meta-Analyses Extension for Scoping Reviews to improve transparency and repeatability. Two independent evaluators (Capobianco M and Visalli F) performed the database searches and preliminary screening of titles and abstracts. Discrepancies concerning inclusion were addressed through assessment by a third reviewer (Zeppieri M) to mitigate selection bias and guarantee consensus-based literature inclusion.

The synthesis was conducted qualitatively, with themes arising from iterative readings of the chosen literature. Themes encompassed diagnostic performance, data quality, generalizability, transparency, hurdles to clinical use, regulatory challenges, and ethical considerations. This review prioritizes narrative interpretation and the integration of data across various study designs, rather than consolidating quantitative indicators like pooled sensitivities or specificities, in alignment with its editorial intent.

CLINICAL APPLICATIONS

AI in ophthalmology has been mostly utilized for illnesses that are the foremost sources of global visual impairment, such as DR, glaucoma, and AMD. These disorders exemplify critical case studies for comprehending both the potential and the drawbacks of AI implementation in clinical care.

DR

DR is the most advanced and successful application of AI among all ocular illnesses. The condition is very suitable for automated screening because of its dependence on fundus photography, the relative homogeneity of diagnostic characteristics, and the urgent necessity for early detection in at-risk groups. Initial deep learning algorithms exhibited sensitivities and specificities akin to those of trained ophthalmologists in detecting referable DR[8,9].

These findings facilitated the development of the inaugural FDA-approved autonomous diagnostic system, IDx-DR, engineered to identify more-than-mild DR without necessitating clinical intervention[10]. Since then, additional autonomous systems have received United States FDA clearances, broadening device and workflow options. Notably, EyeArt obtained clearance for the autonomous detection of both more-than-mild and vision-threatening DR, with subsequent indications expanding supported cameras; more recently, a handheld camera + AI was cleared, pointing toward portable, point-of-care screening[11,12]. These regulatory milestones expand the potential reach of DR screening into primary care, retail clinics, and community venues.

There is also emerging evidence of implementation benefits. Real-world deployments report improved screening completion/adherence when autonomous AI is embedded in primary care workflows, and pragmatic evaluations describe high imageability with practical operator training. While these studies primarily track process metrics (exam completion, imageability, referral yield), they demonstrate that integrating AI can close screening gaps at scale and

reduce specialist workload[13]. Economic evaluations are beginning to quantify value: Model-based analyses suggest DR AI screening can be cost-effective or even cost-saving under realistic sensitivity/specificity and per-screen costs, especially when it increases adherence among underserved populations. That said, cost-effectiveness depends on local prices, adherence baselines, and referral capacity, so generalizability should be assessed site-by-site[14]. However, the transition from proof-of-concept to broad deployment has revealed substantive obstacles and mixed results. For example, a multicenter head-to-head validation study of seven AI DR tools found wide variation in real-world performance - sensitivities ranged from approximately 51% to approximately 86% and specificities from approximately 60% to approximately 84%, highlighting the challenge of maintaining performance across diverse populations and imaging conditions [15].

Empirical studies have underscored significant challenges in scaling AI systems for DR. In community or real-world screening programs, inconsistent image quality (due to varying cameras, lighting, and operator technique) may lead to variable algorithmic performance[16]. Furthermore, many AI-based DR tools have been developed in isolation from comprehensive diabetes care pathways; the lack of integration with systemic diabetes management and follow-up limits their potential to meaningfully prevent vision loss or improve long-term patient outcomes[17]. Moreover, barriers to integration remain substantial. A study using telemedicine and AI DR screening programs flagged cost, infrastructure (information technology and connectivity), workflow redesign, training of personnel, and ensuring return on investment as hurdles[18].

While process outcomes are encouraging, the field still needs more prospective trials powered for patient-level endpoints (vision preservation, treatment timeliness, quality of life) and equity audits across demographic subgroups. Current evidence supports feasibility, accuracy, and workflow gains; the next step is linking those gains to reductions in vision loss and durable health-economic benefit.

Glaucoma

Glaucoma constitutes a more intricate application for AI due to its varied clinical manifestations and the multiple aspects of disease progression. Algorithms have been developed to identify glaucomatous optic neuropathy from fundus images and to measure retinal nerve fiber layer thickness utilizing OCT. Deep-learning systems trained on color fundus photographs can detect glaucomatous optic neuropathy with expert-level performance in clinical datasets (sensitivity/specificity > 90%), though sensitivity drops on multiethnic/external sets - underscoring generalizability challenges[19]. A recent meta-analysis confirms high pooled diagnostic accuracy for glaucoma detection using both fundus photos and OCT[20]. On the structural side, OCT-based models have achieved strong performance for early glaucoma detection and characterization: Ran *et al*[21] trained and validated a deep-learning system on spectral-domain OCT and showed accurate identification of glaucomatous optic neuropathy. Complementing this, Mariottoni *et al*[22] demonstrated “structure-to-function” mapping with AI, learning direct correspondences between OCT-derived measures and visual field loss - an approach that moves beyond binary classification toward patient-specific functional inference from imaging. On the functional side, AI is proving useful for analyzing and forecasting visual field behavior, which is central to clinical decision-making. Asaoka *et al*[23] showed that neural networks trained on standard automated perimetry can detect pre-perimetric glaucoma with high accuracy and predict subsequent damage trajectories, suggesting that learned patterns in visual field data can flag risk before conventional criteria are met. Extending the time horizon, Wen *et al*[24] used deep learning to forecast future Humphrey visual fields up to five years ahead, illustrating the feasibility of long-range functional prediction that could tailor monitoring intervals and treatment intensity to individual risk.

Recent research has concentrated on the integration of structural and functional factors, amalgamating OCT and visual field data to forecast disease progression. Despite their technical sophistication, these models are predominantly limited to research settings, and their efficacy in postponing blindness has not been substantiated in prospective studies[25]. Trust and interpretability continue to pose significant obstacles in glaucoma, as doctors frequently necessitate an elucidation of the methodologies employed by AI algorithms in generating risk predictions prior to their integration into long-term treatment strategies. Taken together, these strands indicate a pathway toward integrated glaucoma AI: OCT-based models for early structural change, structure-function mappings to bridge anatomy and performance, and visual field-based predictors to anticipate progression. In practice, such tools could help triage high-risk patients, optimize follow-up cadence, and guide escalation decisions. Yet, as with other ophthalmic AI, translation will hinge on robust external validation across devices and populations, calibration monitoring over time, and prospective studies showing that earlier detection or improved prediction changes management and ultimately slows vision loss.

AMD and diabetic macular edema

AI development for AMD has focused on three complementary fronts: (1) Image-based disease detection and activity quantification, primarily from OCT; (2) Risk prediction for conversion and progression (to exudation or geographic atrophy); and (3) Treatment guidance (*e.g.*, forecasting anti-vascular endothelial growth factor demand or optimal retreatment intervals). Deep-learning systems reliably segment and quantify hallmark OCT features - intraretinal and subretinal fluid, pigment epithelial detachment, and other lesion components - with agreement to reading-center grades, enabling automated disease-activity metrics at scale. Landmark work demonstrated fully automated, validated fluid detection/quantification on clinical OCT, with subsequent studies generalizing to multi-feature segmentation across neovascular AMD and atrophic AMD[26-28]. Beyond structural OCT, AI models using OCT angiography can diagnose and segment macular neovascularization and explore disease activity from vascular patterns - an emerging avenue that complements fluid-based activity signals[29]. In a notable study using three-dimensional-OCT, investigators integrated automated tissue and pathology segmentations with a risk model to predict fellow-eye conversion to exudative AMD[30]. In another study, using color fundus photography at the population scale, models trained on Age-Related Eye Disease Study/Age-Related Eye Disease Study 2 cohorts (optionally combined with genotype data) estimate the probability and

timing of progression to late AMD, supporting risk-stratified follow-up and trial enrichment[31]. For geographic atrophy, deep learning predicts short-term conversion from intermediate AMD and estimates individual lesion growth from a single baseline OCT or fundus autofluorescence/OCT inputs, opening paths to personalized surveillance and therapy selection[32,33]. Building on automated fluid/feature quantification, machine-learning and deep-learning models predict anti-vascular endothelial growth factor treatment need, retreatment intervals, or 2-year demand using early-phase imaging/clinical data - aimed at supporting treat-and-extend decisions and reducing visit burden[34,35]. A meaningful implementation milestone is the FDA *De Novo* authorization of the Notal Vision SCANLY® Home OCT (2024), a patient-operated device with an AI-based analyzer for at-home visualization/assessment of intraretinal and subretinal fluid between clinic visits. Importantly, its labeling clarifies it is not intended to make treatment decisions or replace in-office care, underscoring both the potential and the current guardrails for AI-enabled home monitoring[36].

As with other ophthalmic AI, domain shift - across OCT vendors (spectral-domain-OCT *vs* swept-source) and acquisition protocols can affect calibration and sensitivity/specificity. OCT angiography-based activity detection adds further variability related to artifacts and segmentation. Robust external validation is therefore crucial before routine adoption. Finally, while AI-derived biomarkers and forecasts are increasingly accurate, prospective trials powered for patient-level endpoints (vision preservation, treatment timeliness, quality of life, cost-effectiveness) remain sparse relative to the volume of development studies; closing this gap is essential to translate pixel-level gains into sustained clinical benefit[37,38].

CURRENT LIMITATIONS

Translating AI from pixel-level efficacy to patient-level advantage in ophthalmology is hindered by a complex array of methodological, clinical, and technical obstacles. The primary problem is the quality and representativeness of the data. Models are typically trained on meticulously curated image datasets that significantly differ from standard clinical data in terms of image quality, disease spectrum, and demographic composition, resulting in overly optimistic accuracy estimates and inadequate generalizability when implemented in novel environments or on different devices[39].

Ground-truth integrity signifies a secondary, often overlooked constraint. Numerous ophthalmic AI systems rely on unreliable reference standards, such as single-grader annotations, heterogeneous clinical records, or flawed diagnostic proxies. Inter-grader heterogeneity in fundus and OCT interpretation, recognized by clinicians, contributes to the labels used for training models, rendering “label noise” a fundamental characteristic of the dataset rather than a mere random disturbance. Inconsistent adjudication methods across several sites and studies generate concealed heterogeneity that diminishes perceived repeatability and amplifies confidence intervals when models are applied outside the laboratory. When ground truth is established by subsequent judgments (*e.g.*, administered treatment), models may learn policy instead of pathology, encoding localized practice patterns as if they were universal clinical truths. The result is shortcut learning, wherein networks leverage misleading correlations - such as device signatures, clinic-specific artifacts, or demographic proxies - rather than disease phenotypes, leading to fragile systems that falter when superficial indicators vanish. Systematic methods for label curation, such as multi-reader adjudication, consensus processes, and uncertainty-aware training, are uncommon, hence constraining the reproducibility and transportability of the presented findings.

Another constraint pertains to robustness and drift. Despite initial validation being robust, performance may deteriorate over time due to advancements in imaging hardware, changes in acquisition methods, or variations in patient demographics - typical instances of covariate and concept drift that are seldom routinely evaluated in ophthalmology. The opaque characteristics of deep learning exacerbate this danger, as failure mechanisms are challenging to predict and comprehend. At the same time, confidence scores serve as inadequate indicators of reliability at the individual case level. Safety-critical scenarios - such as autonomous triage for referable DR or treatment-guiding OCT analysis in neovascular AMD - intensify the repercussions of false negatives and mis-calibrated thresholds; however, few ophthalmic AI studies offer prospective drift surveillance plans or pre-defined recalibration strategies. Out-of-distribution detection - identifying images beyond the model’s training manifold - remains underdeveloped in clinical workflows, permitting unnoticed performance degradation on unusual diseases, atypical phenotypes, or non-standard acquisition conditions [40]. These technological deficiencies immediately erode physician trust and generate medicolegal risks that hinder adoption, despite seemingly robust headline AUCs in retrospective evaluations.

The ability to explain and human-AI interaction constitutes a fourth restriction. In practice, saliency maps and other post hoc explanations sometimes emphasize areas that do not align with established pathophysiology, and they may remain consistent even when model projections change, thereby eroding clinician confidence during decision-making. Furthermore, explanations are seldom assessed as interventions; it is infrequent to examine if the inclusion of an explanation enhances clinician accuracy, calibration, or efficiency in a randomized, workflow-integrated study. For example, in glaucoma, risk predictions that combine structural and functional data may be clinically attractive; however, without a clear rationale for the predicted trajectory and evidence of enhanced progression management, clinicians find it challenging to integrate them into longitudinal care plans.

Regulatory, reporting, and trial design deficiencies further constrain translation. A significant portion of the ophthalmic AI research is retrospective, confined to single centers, and specialized to particular devices, exhibiting a scarcity of prespecified analysis strategies, decision thresholds, or patient-centered objectives[6]. Regulatory approvals for specific tasks, such as autonomous DR diagnosis, have shown feasibility[10]; however, comprehensive frameworks for adaptive learning systems, clinically significant change measures, and ongoing performance evaluation are still in development within ophthalmology[40]. In the absence of standardized reporting protocols and deployment trials that reflect real-world variability, singular measurements like area under the curve will persist in exaggerating clinical significance and

minimizing danger.

Concerns regarding ethics and equity are widespread. The underrepresentation of specific ancestries, age demographics, or comorbidity profiles in training datasets might introduce structural bias into model performance, resulting in consistently inferior sensitivity or specificity for such groups at the same decision thresholds. Even when average performance is satisfactory, uneven calibration among subpopulations might skew triage priorities and resource distribution, prompting equity concerns in screening initiatives and treatment waiting lists. Ultimately, responsibility in shared-control circumstances remains murky; when doctors either overrule or accept an autonomous decision, the routes to accountability are unclear, hindering adoption in risk-averse environments despite significant potential advantages [41].

Barriers to economic and workflow integration finalize the scenario. Many AI systems in ophthalmology require standardized image acquisition and integration with picture archiving and communication systems or electronic health records to ensure seamless interoperability across devices and sites. In parallel, they often depend on high-performance computing infrastructure, including servers equipped with graphics processing units, to enable rapid image processing and model inference in real time[42]. The infrequent inclusion of these expenses in development papers, combined with the lack of comprehensive cost-effectiveness evaluations, means that health systems often lack a compelling business case for adoption, especially when the superiority of clinical outcomes remains unverified. Incorporating a model into busy clinics is complex; it may lead to delays, new points of failure, or redundant tasks if the results are deemed unreliable or cannot be integrated into current care pathways. In screening scenarios, referral networks must be equipped to accommodate heightened case detection without generating bottlenecks; otherwise, improvements in sensitivity at the initial stage result in prolonged wait times and unintended detriment for patients with urgent conditions. Ongoing advantages necessitate persistent monitoring dashboards, management of alert fatigue, and explicit escalation protocols - capabilities that the majority of ophthalmic practices already lack.

Safety hazards remain prevalent despite elevated average performance. False negatives in autonomous DR screening can postpone referrals for sight-threatening conditions. At the same time, erroneous segmentation of OCT layers in macular disorders can misguide treatment time, especially in clinics dependent on a singular imaging modality for decision-making[40]. Calibration drift can transform an initially advantageous threshold into a detrimental one as case mix changes, and few implementations incorporate automatic recalibration methods linked to anticipated ground-truth sampling or clinician input[43].

The limitations mentioned above elucidate why pixel-perfect classification has not produced consistent, quantifiable enhancements in visual results, cost efficiency, or equality in glaucoma, diabetic eye disease, and macular diseases. The way forward is not to forsake AI, but to redirect the discipline from performance measurements to clinical validity, safety, and sustainability. This entails prospectively establishing decision criteria associated with actions, checking subgroup performance and calibrating over time, incorporating human-factors design into interfaces and processes, and assessing impact using outcomes significant to patients and health systems. Unless studies are designed and presented with this translational objective, AI in ophthalmology will continue to produce remarkable visuals without ensuring benefits for individuals. **Figure 1** summarizes the limits of using AI in managing ophthalmologic disorders.

FUTURE OUTLOOKS

The constraints of AI in ophthalmology underscore the need to shift research focus from proof-of-concept demonstrations to clinically verified applications. While algorithms have reliably attained notable diagnostic precision in retrospective investigations, the subsequent stage of advancement must exhibit significant effects on patient outcomes, workflow efficiency, and health equity. This necessitates a transition from strictly defined performance indicators to a comprehensive, patient-centered assessment. A critical area for future research is the execution of stringent prospective studies. The bulk of published research is retrospective, single-center, and concentrates on highly curated datasets that significantly differ from the variability observed in real-world practice. Future multicenter studies are required to assess algorithms in varied populations, employing pragmatic designs that consider comorbidities, heterogeneity in image quality, and varying healthcare infrastructures. Such trials must have patient-centered objectives, including the preservation of visual function, a reduction in time to diagnosis, enhancement of quality of life, and evident cost-effectiveness. Randomized designs, stepped-wedge cluster studies, and post-deployment surveillance frameworks can ensure that algorithms maintain validity post-deployment and adapt securely to changing patient populations and advancing technologies.

A further focal point is transparency and elucidation. For significant adoption of AI in ophthalmology, physicians must comprehend not just the system's predictions but also the rationale behind such predictions. Progress in explainable AI, including saliency mapping, attention-based visualization, and uncertainty quantification, is promising; nonetheless, it requires thorough validation in clinical trials. Generating visually appealing explanations is inadequate; these tools must enhance clinician trust, calibration, and decision-making precision in practical workflows. Human-factors engineering is crucial for designing interfaces that interact effortlessly with electronic health records and imaging systems, reduce cognitive load, and enhance rather than supplant physician skills.

Equity continues to be an essential factor for forthcoming research. A significant number of datasets utilized for AI research are predominantly sourced from high-resource areas, which can jeopardize the performance for underrepresented people. This disparity threatens to exacerbate systemic imbalances in eye care, especially in low-income nations where the incidence of preventable blindness is most pronounced. Incorporating global populations into datasets, implementing learning methodologies that enable distributed model training without direct data exchange, and con-

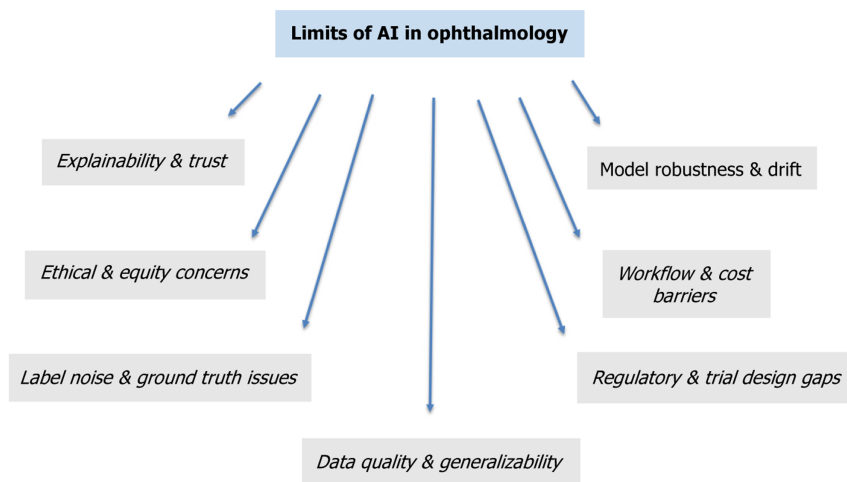


Figure 1 Flowchart delineating the principal constraints of artificial intelligence in the management of ophthalmic conditions. Primary obstacles encompass data quality and generalizability, label noise and inconsistencies in ground truth, robustness and drift of deployed models, insufficient ability to explain, clinician trust, deficiencies in regulatory frameworks and clinical trial design, ethical and equity issues, challenges in workflow integration, and costs. These aspects elucidate why elevated image-level accuracy does not necessarily result in enhanced patient outcomes. AI: Artificial intelligence.

ducting regular equity audits across demographic subgroups are essential measures to ensure the equitable distribution of AI benefits. Explicitly addressing prejudice is crucial for AI to fulfill its potential as a democratizing influence in ophthalmology, rather than a technology that exacerbates disparities in access and outcomes.

In the future, the discipline must adopt adaptable and ever-evolving systems. Static algorithms established upon deployment are inadequate for dynamic clinical settings, as case mix, imaging technologies, and treatment methodologies change over time. Implementing safeguards for continuous learning, such as version control, out-of-distribution detection, and uncertainty-aware recalibration, can guarantee stability while facilitating adaptability. These techniques must be complemented with monitoring dashboards that assess subgroup performance, safety signals, and model drift in real time, thus allowing physicians and regulators to respond proactively when performance diverges.

The amalgamation of multimodal data signifies a promising horizon. Instead of limiting AI to unimodal image categorization, next-generation systems are expected to integrate imaging with genetic, systemic, and functional data streams to construct comprehensive models of illness risk and progression. In glaucoma, integrating OCT measures, visual field data, systemic risk factors, and genetic predispositions may yield individualized progression predictions that inform monitoring intervals and therapeutic approaches. In DR, correlating retinal imaging with systemic indicators such as glycemic regulation, renal function, and cardiovascular health may facilitate comprehensive risk assessment that extends beyond ocular considerations. This multimodal integration advances AI towards the goal of precision medicine, necessitating sophisticated data harmonization pipelines and stringent patient privacy protections.

Regulatory frameworks and reporting standards must advance concurrently with technology developments. Frameworks like Consolidated Standards of Reporting Trials-AI and Standard Protocol Items: Recommendations for Interventional Trials-AI offer a foundational reference[44]. Future ophthalmic AI research should implement pre-registration, predetermined thresholds, consistent calibration reporting, and transparent subgroup analyses as essential criteria. Regulators must confront the problems presented by adaptive systems that evolve over time, necessitating flexible approval processes that reconcile innovation with patient safety. Cost-effectiveness assessments, health economic modeling, and public reporting of implementation expenses will enhance the justification for adoption by healthcare systems and payers.

The future of AI in ophthalmology will hinge on collaborative ecosystems uniting clinicians, data scientists, patients, ethicists, and regulators. Algorithms can only be built to accord with clinical reality, uphold ethical standards, and yield measurable enhancements in patient outcomes through ongoing debate and collaborative development. The forthcoming decade will certainly ascertain whether AI evolves into a fundamental element of precision ophthalmology or persists as an outstanding yet underexploited technological artifact.

CONCLUSION

AI has swiftly emerged as one of the most promising technological advancements in ophthalmology. Its capacity to evaluate extensive imaging data, identify nuanced patterns, and achieve or exceed expert-level diagnostic precision in controlled environments has justifiably garnered international interest. Nevertheless, despite these accomplishments, the translation of algorithmic efficacy into quantifiable enhancements for patients has remained constrained. The disparity between pixel-level accuracy and clinical results remains the primary difficulty in this domain. Current evidence indicates that AI systems in ophthalmology are promising but encounter enduring obstacles: Data bias, label noise, spectrum effects, and inadequate generalizability compromise reliability beyond controlled environments. The opacity of

black-box systems undermines physician trust, and the lack of prospective, randomized trials limits confidence in their real-world applicability. Regulatory ambiguity, elevated integration costs, and insufficient focus on patient-centered outcomes impede implementation. Ethical considerations - specifically justice, fairness, and accountability - exacerbate these challenges, increasing the likelihood that AI may unintentionally exacerbate gaps in eye care. Nonetheless, the trend of the study indicates a positive direction ahead. Stringent external validation, comprehensive prospective trials, and ongoing monitoring are crucial for guaranteeing safe and effective implementation. The degree of ability to explain, transparency, and the design of human-AI interactions will influence physicians' acceptance or rejection of these tools. Creating inclusive global datasets, utilizing federated learning, and doing equity audits will ensure that AI serves all patients fairly. Multimodal integration has the potential to advance AI from mere picture classification to precision medicine, while adaptive systems with post-deployment recalibration will provide sustained reliability over time. The development of regulatory frameworks and reporting standards must parallel these scientific advancements to establish an ecosystem of accountability, repeatability, and clinical trust. The essential message for doctors and researchers is unequivocal: AI in ophthalmology should be evaluated not solely on its ability to classify images in retrospective datasets, but on its efficacy in enhancing patient outcomes when integrated into clinical processes. AI can only realize its transformative potential by advancing beyond mere technical performance to encompass clinical validation, equity, and patient-centered benefits. For ophthalmologists, the priority is not the decision to adopt AI, but rather how to guarantee its implementation is safe, transparent, equitable, and therapeutically significant.

REFERENCES

- 1 **Ting DSW**, Pasquale LR, Peng L, Campbell JP, Lee AY, Raman R, Tan GSW, Schmetterer L, Keane PA, Wong TY. Artificial intelligence and deep learning in ophthalmology. *Br J Ophthalmol* 2019; **103**: 167-175 [RCA] [PMID: 30361278 DOI: 10.1136/bjophthalmol-2018-313173] [FullText] [Full Text(PDF)]
- 2 **Khan Z**, Gaidhane AM, Singh M, Ganesan S, Kaur M, Sharma GC, Rani P, Sharma R, Thapliyal S, Kushwaha M, Kumar H, Agarwal RK, Shabil M, Verma L, Sidhu A, Manan NBA, Bushi G, Mehta R, Sah S, Satapathy P, Samal SK. Diagnostic Accuracy of IDX-DR for Detecting Diabetic Retinopathy: A Systematic Review and Meta-Analysis. *Am J Ophthalmol* 2025; **273**: 192-204 [RCA] [PMID: 39986640 DOI: 10.1016/j.ajo.2025.02.022] [FullText]
- 3 **Labib KM**, Ghumman H, Jain S, Jarstad JS. A Review of the Utility and Limitations of Artificial Intelligence in Retinal Disorders and Pediatric Ophthalmology. *Cureus* 2024; **16**: e71063 [RCA] [PMID: 39380780 DOI: 10.7759/cureus.71063] [FullText]
- 4 **Tseng RMWW**, Gunasekaran DV, Tan SSH, Rim TH, Lum E, Tan GSW, Wong TY, Tham YC. Considerations for Artificial Intelligence Real-World Implementation in Ophthalmology: Providers' and Patients' Perspectives. *Asia Pac J Ophthalmol (Phila)* 2021; **10**: 299-306 [RCA] [PMID: 34383721 DOI: 10.1097/APO.0000000000000400] [FullText]
- 5 **Chen S**, Bai W. Artificial intelligence technology in ophthalmology public health: current applications and future directions. *Front Cell Dev Biol* 2025; **13**: 1576465 [RCA] [PMID: 40313720 DOI: 10.3389/fcell.2025.1576465] [FullText] [Full Text(PDF)]
- 6 **Li Z**, Wang L, Wu X, Jiang J, Qiang W, Xie H, Zhou H, Wu S, Shao Y, Chen W. Artificial intelligence in ophthalmology: The path to the real-world clinic. *Cell Rep Med* 2023; **4**: 101095 [RCA] [PMID: 37385253 DOI: 10.1016/j.xcrm.2023.101095] [FullText] [Full Text(PDF)]
- 7 **Nakayama LF**, Matos J, Quion J, Novaes F, Mitchell WG, Mwavu R, Hung CJJ, Santiago APD, Phanphruk W, Cardoso JS, Celi LA. Unmasking biases and navigating pitfalls in the ophthalmic artificial intelligence lifecycle: A narrative review. *PLOS Digit Health* 2024; **3**: e0000618 [RCA] [PMID: 39378192 DOI: 10.1371/journal.pdig.0000618] [FullText]
- 8 **Abramoff MD**, Lou Y, Erginay A, Clarida W, Amelon R, Folk JC, Niemeijer M. Improved Automated Detection of Diabetic Retinopathy on a Publicly Available Dataset Through Integration of Deep Learning. *Invest Ophthalmol Vis Sci* 2016; **57**: 5200-5206 [RCA] [PMID: 27701631 DOI: 10.1167/iovs.16-19964] [FullText]
- 9 **Gulshan V**, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, Venugopalan S, Widner K, Madams T, Cuadros J, Kim R, Raman R, Nelson PC, Mega JL, Webster DR. Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs. *JAMA* 2016; **316**: 2402-2410 [RCA] [PMID: 27898976 DOI: 10.1001/jama.2016.17216] [FullText]
- 10 **van der Heijden AA**, Abramoff MD, Verbraak F, van Hecke MV, Liem A, Nijpels G. Validation of automated screening for referable diabetic retinopathy with the IDx-DR device in the Hoorn Diabetes Care System. *Acta Ophthalmol* 2018; **96**: 63-68 [RCA] [PMID: 29178249 DOI: 10.1111/aos.13613] [FullText] [Full Text(PDF)]
- 11 **Vujosevic S**, Limoli C, Nucci P. Novel artificial intelligence for diabetic retinopathy and diabetic macular edema: what is new in 2024? *Curr Opin Ophthalmol* 2024; **35**: 472-479 [RCA] [PMID: 39259647 DOI: 10.1097/ICU.0000000000001084] [FullText]
- 12 **Lim JJ**, Regillo CD, Sadda SR, Ipp E, Bhaskaranand M, Ramachandra C, Solanki K. Artificial Intelligence Detection of Diabetic Retinopathy: Subgroup Comparison of the EyeArt System with Ophthalmologists' Dilated Examinations. *Ophthalmol Sci* 2023; **3**: 100228 [RCA] [PMID: 36345378 DOI: 10.1016/j.xops.2022.100228] [FullText] [Full Text(PDF)]
- 13 **Duggal M**, Chauhan A, Gupta V, Kankaria A, Budhija D, Verma P, Miglani V, Syal P, Kaur G, Kumar L, Mutyala N, Bezbaruah R, Sood N, Kernohan A, Menon G, Vale L. Real-World Evaluation of AI-Driven Diabetic Retinopathy Screening in Public Health Settings: Validation and Implementation Study. *JMIR Med Inform* 2025; **13**: e67529 [RCA] [PMID: 40925861 DOI: 10.2196/67529] [FullText] [Full Text(PDF)]
- 14 **Wu H**, Jin K, Yip CC, Koh V, Ye J. A systematic review of economic evaluation of artificial intelligence-based screening for eye diseases: From possibility to reality. *Surv Ophthalmol* 2024; **69**: 499-507 [RCA] [PMID: 38492584 DOI: 10.1016/j.survophthal.2024.03.008] [FullText]
- 15 **Lee AY**, Yanagihara RT, Lee CS, Blazes M, Jung HC, Chee YE, Gencarella MD, Gee H, Maa AY, Cockerham GC, Lynch M, Boyko EJ. Multicenter, Head-to-Head, Real-World Validation Study of Seven Automated Artificial Intelligence Diabetic Retinopathy Screening Systems. *Diabetes Care* 2021; **44**: 1168-1175 [RCA] [PMID: 33402366 DOI: 10.2337/dc20-1877] [FullText] [Full Text(PDF)]
- 16 **Chawla R**, Karkhanis P, Shah M, Das A, Sharma R, Almaula D, Venkatesh P, Singh HV, Kumar M, Samanta R, Kumarl V, Shah A, Vadera B, Jain N, Sen A, Shreedhar S, Garg V, Dhaval S, Ganesh K, Rana S, Tandon R. Artificial intelligence for advancing eye care in resource-poor settings: Assessing the predictive accuracy of an AI-model for diabetic retinopathy screening in India. *Glob Epidemiol* 2025; **9**: 100209 [RCA] [PMID: 40538414 DOI: 10.1016/j.gloepi.2025.100209] [FullText] [Full Text(PDF)]
- 17 **Hassan S**, Ibrahim S, Bielecki J, Stanimirovic A, Mathew S, Hooey R, Bowen JM, Rac VE. Integrating artificial intelligence in community-

- based diabetes care programmes: enhancing inclusiveness, diversity, equity and accessibility a realist review protocol. *BMJ Open* 2025; **15**: e100512 [RCA] [PMID: 40664423 DOI: 10.1136/bmjopen-2025-100512] [FullText] [Full Text(PDF)]
- 18 **Chen JS**, Lin MC, Yiu G, Thorne C, Kulasa K, Stewart J, Nudleman E, Freeby M, Han MA, Baxter SL. Barriers to Implementation of Teleretinal Diabetic Retinopathy Screening Programs Across the University of California. *Telemed J E Health* 2023; **29**: 1810-1818 [RCA] [PMID: 37256712 DOI: 10.1089/tmj.2022.0489] [FullText]
 - 19 **Liu H**, Li L, Wormstone IM, Qiao C, Zhang C, Liu P, Li S, Wang H, Mou D, Pang R, Yang D, Zangwill LM, Moghimi S, Hou H, Bowd C, Jiang L, Chen Y, Hu M, Xu Y, Kang H, Ji X, Chang R, Tham C, Cheung C, Ting DSW, Wong TY, Wang Z, Weinreb RN, Xu M, Wang N. Development and Validation of a Deep Learning System to Detect Glaucomatous Optic Neuropathy Using Fundus Photographs. *JAMA Ophthalmol* 2019; **137**: 1353-1360 [RCA] [PMID: 31513266 DOI: 10.1001/jamaophthalmol.2019.3501] [FullText]
 - 20 **Ling XC**, Chen HS, Yeh PH, Cheng YC, Huang CY, Shen SC, Lee YS. Deep Learning in Glaucoma Detection and Progression Prediction: A Systematic Review and Meta-Analysis. *Biomedicines* 2025; **13**: 420 [RCA] [PMID: 40002833 DOI: 10.3390/biomedicines13020420] [Full Text]
 - 21 **Ran AR**, Cheung CY, Wang X, Chen H, Luo LY, Chan PP, Wong MOM, Chang RT, Mannil SS, Young AL, Yung HW, Pang CP, Heng PA, Tham CC. Detection of glaucomatous optic neuropathy with spectral-domain optical coherence tomography: a retrospective training and validation deep-learning analysis. *Lancet Digit Health* 2019; **1**: e172-e182 [RCA] [PMID: 33323187 DOI: 10.1016/S2589-7500(19)30085-8] [FullText]
 - 22 **Mariottoni EB**, Datta S, Dov D, Jammal AA, Berchuck SI, Tavares IM, Carin L, Medeiros FA. Artificial Intelligence Mapping of Structure to Function in Glaucoma. *Transl Vis Sci Technol* 2020; **9**: 19 [RCA] [PMID: 32818080 DOI: 10.1167/tvst.9.2.19] [FullText] [Full Text(PDF)]
 - 23 **Asaoka R**, Murata H, Hirasawa K, Fujino Y, Matsuura M, Miki A, Kanamoto T, Ikeda Y, Mori K, Iwase A, Shoji N, Inoue K, Yamagami J, Araie M. Using Deep Learning and Transfer Learning to Accurately Diagnose Early-Onset Glaucoma From Macular Optical Coherence Tomography Images. *Am J Ophthalmol* 2019; **198**: 136-145 [RCA] [PMID: 30316669 DOI: 10.1016/j.ajo.2018.10.007] [FullText]
 - 24 **Wen JC**, Lee CS, Keane PA, Xiao S, Rokem AS, Chen PP, Wu Y, Lee AY. Forecasting future Humphrey Visual Fields using deep learning. *PLoS One* 2019; **14**: e0214875 [RCA] [PMID: 30951547 DOI: 10.1371/journal.pone.0214875] [FullText] [Full Text(PDF)]
 - 25 **Tonti E**, Tonti S, Mancini F, Bonini C, Spadea L, D'Esposito F, Gagliano C, Musa M, Zeppieri M. Artificial Intelligence and Advanced Technology in Glaucoma: A Review. *J Pers Med* 2024; **14**: 1062 [RCA] [PMID: 39452568 DOI: 10.3390/jpm14101062] [FullText] [Full Text (PDF)]
 - 26 **Schlegl T**, Waldstein SM, Bogunovic H, Endstraßer F, Sadeghipour A, Philip AM, Podkowinski D, Gerendas BS, Langs G, Schmidt-Erfurth U. Fully Automated Detection and Quantification of Macular Fluid in OCT Using Deep Learning. *Ophthalmology* 2018; **125**: 549-558 [RCA] [PMID: 29224926 DOI: 10.1016/j.ophtha.2017.10.031] [FullText]
 - 27 **Borrelli E**, Oakley JD, Iaccarino G, Russakoff DB, Battista M, Grosso D, Borghesan F, Barresi C, Sacconi R, Bandello F, Querques G. Deep-learning based automated quantification of critical optical coherence tomography features in neovascular age-related macular degeneration. *Eye (Lond)* 2024; **38**: 537-544 [RCA] [PMID: 37670143 DOI: 10.1038/s41433-023-02720-8] [FullText]
 - 28 **Liefers B**, Taylor P, Alsaedi A, Bailey C, Balaskas K, Dhinra N, Egan CA, Rodrigues FG, Gonzalo CG, Heeren TFC, Lotery A, Müller PL, Olvera-Barrios A, Paul B, Schwartz R, Thomas DS, Warwick AN, Tufail A, Sánchez CI. Quantification of Key Retinal Features in Early and Late Age-Related Macular Degeneration Using Deep Learning. *Am J Ophthalmol* 2021; **226**: 1-12 [RCA] [PMID: 33422464 DOI: 10.1016/j.ajo.2020.12.034] [FullText]
 - 29 **Wang J**, Hormel TT, Tsuboi K, Wang X, Ding X, Peng X, Huang D, Bailey ST, Jia Y. Deep Learning for Diagnosing and Segmenting Choroidal Neovascularization in OCT Angiography in a Large Real-World Data Set. *Transl Vis Sci Technol* 2023; **12**: 15 [RCA] [PMID: 37058103 DOI: 10.1167/tvst.12.4.15] [FullText]
 - 30 **Yim J**, Chopra R, Spitz T, Winkens J, Obika A, Kelly C, Askham H, Lukic M, Huemer J, Fasler K, Moraes G, Meyer C, Wilson M, Dixon J, Hughes C, Rees G, Khaw PT, Karthikesalingam A, King D, Hassabis D, Suleyman M, Back T, Ledsam JR, Keane PA, De Fauw J. Predicting conversion to wet age-related macular degeneration using deep learning. *Nat Med* 2020; **26**: 892-899 [RCA] [PMID: 32424211 DOI: 10.1038/s41591-020-0867-7] [FullText]
 - 31 **Peng Y**, Keenan TD, Chen Q, Agrón E, Allot A, Wong WT, Chew EY, Lu Z. Predicting risk of late age-related macular degeneration using deep learning. *NPJ Digit Med* 2020; **3**: 111 [RCA] [PMID: 32904246 DOI: 10.1038/s41746-020-00317-z] [FullText] [Full Text(PDF)]
 - 32 **Dow ER**, Jeong HK, Katz EA, Toth CA, Wang D, Lee T, Kuo D, Allingham MJ, Hadziahmetovic M, Mettu PS, Schuman S, Carin L, Keane PA, Henao R, Lad EM. A Deep-Learning Algorithm to Predict Short-Term Progression to Geographic Atrophy on Spectral-Domain Optical Coherence Tomography. *JAMA Ophthalmol* 2023; **141**: 1052-1061 [RCA] [PMID: 37856139 DOI: 10.1001/jamaophthalmol.2023.4659] [Full Text]
 - 33 **Mai J**, Lachinov D, Reiter GS, Riedl S, Grechenig C, Bogunovic H, Schmidt-Erfurth U. Deep Learning-Based Prediction of Individual Geographic Atrophy Progression from a Single Baseline OCT. *Ophthalmol Sci* 2024; **4**: 100466 [RCA] [PMID: 38591046 DOI: 10.1016/j.xops.2024.100466] [FullText] [Full Text(PDF)]
 - 34 **Han JM**, Han J, Ko J, Jung J, Park JI, Hwang JS, Yoon J, Jung JH, Hwang DD. Anti-VEGF treatment outcome prediction based on optical coherence tomography images in neovascular age-related macular degeneration using a deep neural network. *Sci Rep* 2024; **14**: 28253 [RCA] [PMID: 39548212 DOI: 10.1038/s41598-024-79034-6] [FullText]
 - 35 **Gallardo M**, Munk MR, Kurmann T, De Zanet S, Mosinska A, Karagoz IK, Zinkernagel MS, Wolf S, Sznitman R. Machine Learning Can Predict Anti-VEGF Treatment Demand in a Treat-and-Extend Regimen for Patients with Neovascular AMD, DME, and RVO Associated Macular Edema. *Ophthalmol Retina* 2021; **5**: 604-624 [RCA] [PMID: 33971352 DOI: 10.1016/j.oret.2021.05.002] [FullText]
 - 36 **Schneider EW**, Heier JS, Hlekamp NM, Busquets MA, Wagner AL, Mukkamala SK, Riemann CD, Lee SY, Joondeph BC, Houston SS, Nahen K, Mohan N, Benyamini G. Pivotal Trial Toward Effectiveness of Self-administered OCT in Neovascular Age-related Macular Degeneration. Report 2-Artificial Intelligence Analytics. *Ophthalmol Sci* 2025; **5**: 100662 [RCA] [PMID: 39811265 DOI: 10.1016/j.xops.2024.100662] [FullText]
 - 37 **Borrelli E**, Serafino S, Ricardi F, Coletto A, Neri G, Olivieri C, Ulla L, Foti C, Marolo P, Toro MD, Bandello F, Reibaldi M. Deep Learning in Neovascular Age-Related Macular Degeneration. *Medicina (Kaunas)* 2024; **60**: 990 [RCA] [PMID: 38929607 DOI: 10.3390/medicina60060990] [FullText]
 - 38 **Tombolini B**, Crincoli E, Sacconi R, Battista M, Fantaguzzi F, Servillo A, Bandello F, Querques G. Optical Coherence Tomography Angiography: A 2023 Focused Update on Age-Related Macular Degeneration. *Ophthalmol Ther* 2024; **13**: 449-467 [RCA] [PMID: 38180632 DOI: 10.1007/s40123-023-00870-2] [FullText]
 - 39 **Rashidisabet H**, Sethi A, Jindarak P, Edmonds J, Chan RVP, Leiderman YI, Vajaranant TS, Yi D. Validating the Generalizability of

- Ophthalmic Artificial Intelligence Models on Real-World Clinical Data. *Transl Vis Sci Technol* 2023; **12**: 8 [RCA] [PMID: 37922149 DOI: 10.1167/tvst.12.11.8] [FullText]
- 40 Swaminathan U, Daigavane S. Unveiling the Potential: A Comprehensive Review of Artificial Intelligence Applications in Ophthalmology and Future Prospects. *Cureus* 2024; **16**: e61826 [RCA] [PMID: 38975538 DOI: 10.7759/cureus.61826] [FullText]
- 41 Jassar S, Adams SJ, Zarzeczny A, Burbridge BE. The future of artificial intelligence in medicine: Medical-legal considerations for health leaders. *Health Manage Forum* 2022; **35**: 185-189 [RCA] [PMID: 35354409 DOI: 10.1177/08404704221082069] [FullText] [Full Text(PDF)]
- 42 Leiner T, Bennink E, Mol CP, Kuijf HJ, Veldhuis WB. Bringing AI to the clinic: blueprint for a vendor-neutral AI deployment infrastructure. *Insights Imaging* 2021; **12**: 11 [RCA] [PMID: 33528677 DOI: 10.1186/s13244-020-00931-1] [FullText] [Full Text(PDF)]
- 43 Sahiner B, Chen W, Samala RK, Petrick N. Data drift in medical machine learning: implications and potential remedies. *Br J Radiol* 2023; **96**: 20220878 [RCA] [PMID: 36971405 DOI: 10.1259/bjr.20220878] [FullText]
- 44 Solebo AL, Braithwaite T. Implications of the artificial intelligence extensions to the guidelines for consolidated standards of reporting trials and for standard protocol item recommendations for interventional trials (the CONSORT-AI and SPIRIT-AI extensions). *EClinicalMedicine* 2020; **26**: 100536 [RCA] [PMID: 33089126 DOI: 10.1016/j.eclinm.2020.100536] [FullText] [Full Text(PDF)]

FOOTNOTES

Specialty type: Medical laboratory technology

Country of origin: Italy

Author contributions: Zeppieri M, Capobianco M, Visalli F, Khouyyi M, Musa M, Avitabile A, Leandro I, Giglio R, Tognetto D, Gagliano C, D'Esposito F, and Cappellani F wrote the outline, assisted in the writing and editing of the draft and final paper, making critical revisions of the manuscript and viewing all versions of the manuscript; Zeppieri M, Capobianco M, Visalli F, Khouyyi M, Musa M, Avitabile A, and Cappellani F did the research and writing of the manuscript; Zeppieri M, Capobianco M, Visalli F, Khouyyi M, Musa M, Gagliano C, D'Esposito F, and Cappellani F contributed to the scientific editing; Zeppieri M, Gagliano C, D'Esposito F, Cappellani F were responsible for the conception and design of the study. All authors provided the final approval of the article.

AI contribution statement: ChatGPT (OpenAI, GPT-5.3) and Grammarly were used to assist with summarizing existing literature, addressing issues in the rebuttal, and enhancing the flow and English language quality. No AI-generated images were used.

Conflict-of-interest statement: All the authors report no relevant conflicts of interest for this article.

S-Editor: Zuo Q

L-Editor: A

P-Editor: Zhang YL



Published by **Baishideng Publishing Group Inc**
7041 Koll Center Parkway, Suite 160, Pleasanton, CA 94566, USA

Telephone: +1-925-3991568

E-mail: office@baishideng.com

Help Desk: <https://www.f6publishing.com/helpdesk>

<https://www.wjgnet.com>

