Research Article

Artificial intelligence in the diagnosis of inherited retinal dystrophies: from algorithmic accuracy to an integrated clinical pathway

Andrei Nikolaevich Abramov[1]*, Khedi Arbievna Demisheva[2], Iman Umarovna Turlaeva[2], Polina Sergeevna Shelkunova[3], Samir Sakit Ogly Verdiev[3], Alzhana Israpilovna Kurbanova[3], Bella Ruslanovna Kuzheva[3], Darya Alekseevna Bochanova[3], Luiza Taizmovna Barkenkhoeva[4], Milana Umarovna Abyakarova[4]

[1] Faculty of Medicine, Saratov Medical University "Reaviz", Saratov, Russia
[2] Faculty of Medicine, Astrakhan State Medical University, Astrakhan, Russia
[3] Institute of Clinical Medicine, Saratov State Medical University named after V.I. Razumovsky, Saratov, Russia
[4] Faculty of Medicine, Medical Institute named after A.H.Aushev, Ingush State University, Magas, Republic of Ingushetia, Russia
*** Corresponding author**: Andrei Nikolaevich Abramov, e-mail: ab373@mail.ru, ORCID: 0009-0003-4686-340X

Abstract

Inherited retinal dystrophies (IRDs) are a genetically heterogeneous group of monogenic disorders caused by variants in more than 280 genes and are now a leading cause of blindness in working-age adults in several countries. Access to gene therapy and gene-specific clinical trials requires a molecular diagnosis, yet patients often wait years for one, and next-generation sequencing remains inconclusive in a quarter to nearly half of cases. Deep learning models can now recognise IRDs on optical coherence tomography, fundus autofluorescence and colour fundus photographs, quantify structural biomarkers and rank the most probable causative genes. In a 2026 multicentre randomised trial, the Retina4IRD decision support system increased the top-5 accuracy of specialists' gene predictions from 67.3% to 88.5%. Nevertheless, only 7.2% of ophthalmologists report regular use of artificial intelligence (AI). This review weighs the evidence for the diagnostic performance of AI in IRDs against the barriers to its adoption, including insufficient training, cost, automation bias, limited explainability, unrepresentative training data and population differences in the genetic spectrum. We argue that algorithmic accuracy is no longer the main constraint. Integration into the patient pathway, transferability across populations, regulation of continuously updated models and clinician education have become decisive. We propose an integrated pathway in which AI accelerates referral to specialist centres, enriches phenotyping and increases the yield of genetic testing without replacing sequencing or clinical judgement. Research priorities include prospective multicentre trials in diverse populations, patient-relevant outcomes and federated data consortia.

Keywords: inherited retinal dystrophies; artificial intelligence; deep learning; genetic diagnosis; clinical decision support; innovation adoption

1. Introduction

Inherited retinal dystrophies (IRDs) can no longer be regarded as a niche concern within ophthalmology. In England and Wales, they accounted for 20.2% of blindness certifications among people aged 16-64 years, overtaking diabetic retinopathy as the leading cause for the first time in five decades [1]. The genetic prevalence of autosomal recessive forms alone has been estimated at 1 in 1380 individuals, corresponding to approximately 5.5 million affected people worldwide [2]. Pooled clinical data place the prevalence of all IRDs at about 1 in 3450 [3].

The clinical value of a precise genetic diagnosis increased sharply after the approval of voretigene neparvovec for biallelic *RPE65* variants [4]. Dozens of gene-specific products are now in clinical development [5], and enrolment in these programmes is possible only when the genotype is known. Molecular diagnosis, however, faces three constraints: extreme genetic heterogeneity, with more than 280 implicated genes; a shortage of IRD specialists; and the cost of sequencing. As a result, patients wait an average of 6.4 years between symptom onset and diagnosis [6].

Artificial intelligence (AI) offers a way to shorten this path. The retina is one of the few organs in which the phenotype is directly visible on routine images, and many IRDs produce characteristic patterns on optical coherence tomography (OCT) and fundus autofluorescence (FAF). In 2026, the first multicentre randomised trial of a decision support system for IRDs was published: AI assistance raised the accuracy of specialists' gene predictions from 67.3% to 88.5% [7]. At the same time, a global survey found that 69.5% of ophthalmologists recognised the potential of AI, yet only 7.2% used it regularly [8].

This gap between demonstrated efficacy and real-world use is the subject of the present review. We consider AI in IRDs not as an algorithmic innovation but as an organisational one. The review has three aims: (1) to summarise the evidence on the diagnostic capabilities of AI in IRDs; (2) to analyse the barriers to implementation, both those common to medical AI and those specific to rare diseases; and (3) to propose an integrated patient pathway together with the system-level conditions required to implement it.

2. Genetic architecture of IRDs and the diagnostic bottleneck

2.1. Heterogeneity and concentration

IRDs encompass retinitis pigmentosa, Stargardt disease, cone-rod dystrophies, Leber congenital amaurosis, choroideremia and dozens of rarer conditions. The contribution of individual genes is highly uneven. In the largest molecularly characterised cohort to date (3195 families, United Kingdom), causative variants were identified in 135 genes, yet 20 genes accounted for 71.8% of families [9]. The most frequent were *ABCA4* (20.8%), *USH2A* (9.1%), *RPGR* (5.1%), *PRPH2* (4.6%) and *BEST1* (3.9%) [9]. A detailed allelic analysis of the same cohort showed that a small number of recurrent variants account for a substantial share of the disease burden [10].

This structure matters for diagnosis. The long tail of rare genes makes it impossible for any single clinician to know every phenotype. Conversely, the concentration of the burden in a few genes explains why an algorithm trained on a few dozen classes can cover most patients.

2.2. Population specificity of the gene spectrum

The gene spectrum differs substantially between populations. In a Japanese cohort of 1210 pedigrees, the leading genes were *EYS* and *RP1* rather than *ABCA4* [11]. Common Japanese *EYS* variants reach allele frequencies of 2-3% in the local population yet are almost absent from European databases [12]. In Chinese cohorts, *CYP4V2* and population-specific variants in genes for autosomal dominant retinitis pigmentosa rank among the leading causes [9,13]. The implication for AI is direct: the class distribution on which a model is trained is tied to the population that supplied the data (Figure 1).

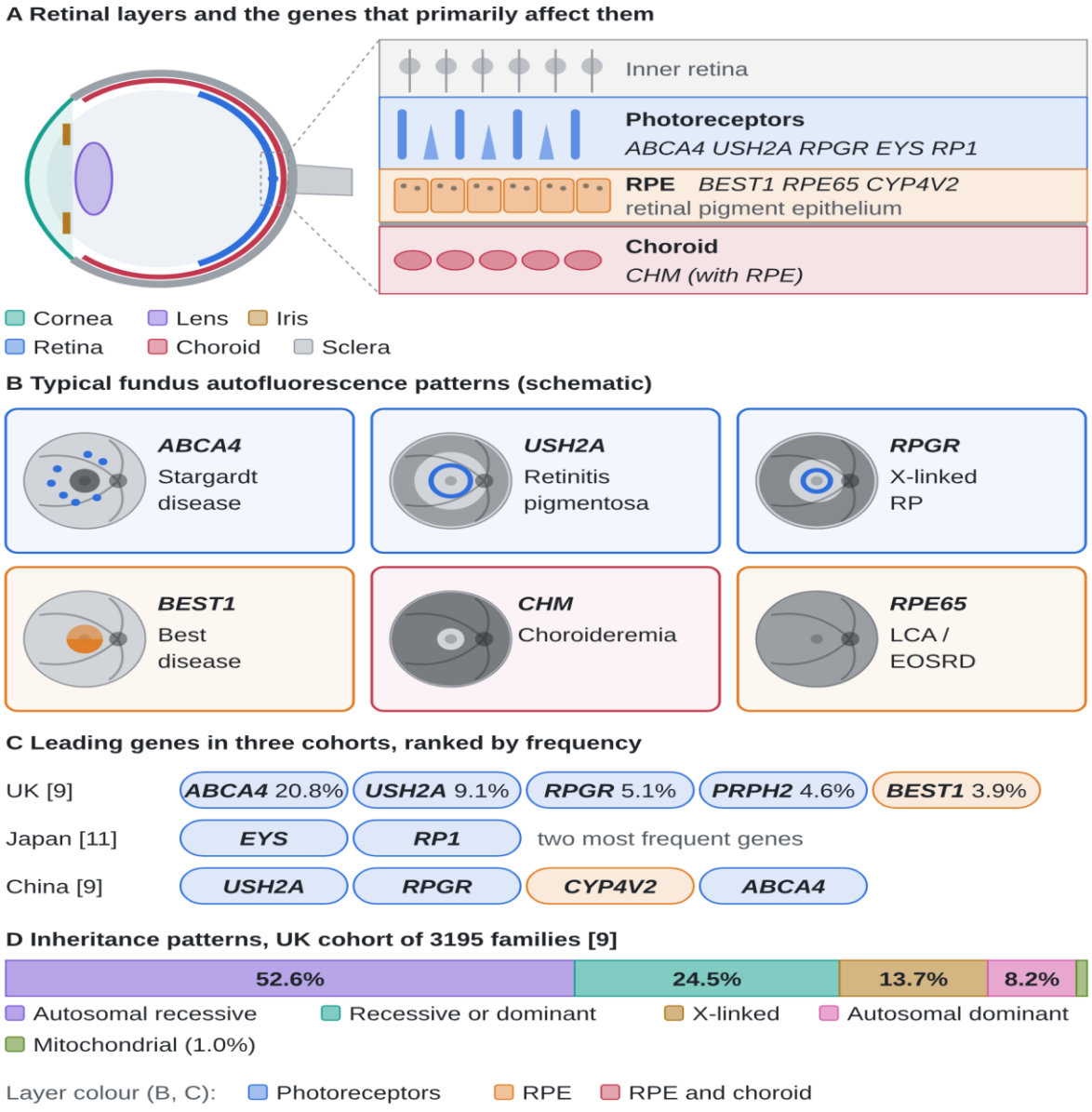

Figure 1. Genetic architecture of inherited retinal dystrophies. (A) Eye and retinal layers, with representative genes grouped by the layer they primarily affect. (B) Schematic fundus autofluorescence patterns of six

representative genes; frame colour indicates the primarily affected layer. (C) Leading genes in UK, Japanese and Chinese cohorts, ranked by frequency, with shares shown where reported. (D) Distribution of inheritance patterns in the UK cohort of 3195 molecularly diagnosed families. Original illustration by the authors based on data from [9,11]. EOSRD, early-onset severe retinal dystrophy; LCA, Leber congenital amaurosis; RP, retinitis pigmentosa; RPE, retinal pigment epithelium.

2.3. The diagnostic bottleneck

Next-generation sequencing has solved the throughput problem of genetic testing but not the problem of diagnosis. A meta-analysis of 105 studies from 28 countries reported a diagnostic yield of 52-74% [14], higher with exome sequencing and with application of the American College of Medical Genetics and Genomics criteria [14]. Yields are lower, and inconclusive results more frequent, in East Asian centres than in European ones [15]. Unsolved cases are not hopeless: targeted, clinician-led reanalysis of exome data increases the number of established diagnoses [16].

The second component of the bottleneck is the time to testing. In a survey of more than 1000 patients from the United Kingdom and Latin America, the mean interval from first symptoms to diagnosis was 6.4 years, and children were more likely than adults to wait longer than a year [6]. Among patients already referred to an ophthalmic genetics service, the picture was worse: 51% had an incorrect initial clinical diagnosis, and the mean delay to the correct diagnosis reached 15 years [17]. The difference reflects the composition of the samples, because specialist centres disproportionately receive patients with the most complex and protracted diagnostic journeys. Most of these referrals come from ophthalmologists [17]. The key point of intervention is therefore the recognition of IRDs in general ophthalmology.

2.4. Why genotype has become therapeutically relevant

The effect of voretigene neparvovec was maintained for 3-4 years in the phase 3 trial [18]. The PERCEIVE registry confirmed its effectiveness in routine practice but also documented chorioretinal atrophy in 12.6% of patients [19]. The number of phase 2-3 gene therapy trials continues to grow; by mid-2024 there were 24, targeting genes from *ABCA4* and *USH2A* to *RPGR* and *CEP290* [20]. For advanced disease, mutation-independent approaches are under development, notably optogenetics, which partially restored visual function in a blind patient with retinitis pigmentosa [21]. Finally, the burden of IRDs is not only clinical: economic losses estimated in Singapore, Europe and North America reach tens of thousands of US dollars per patient per year [22-24].

3. What AI can do: the evidence

Research on AI in IRDs has progressed from binary recognition of a single disease to multiclass prediction of genotype and quantitative assessment of progression. An early systematic review covering studies published up to 2021 noted that such work was scarce and dominated by small single-centre samples [25]. Over the past five years the situation has changed qualitatively, with models trained on tens of thousands of images, external validation across several continents and the first randomised trial (Table 1).

3.1. Recognition of IRDs and differentiation of phenotypes

A meta-analysis of 22 studies demonstrated high diagnostic performance of deep learning models; for Stargardt disease, pooled sensitivity was 96% and pooled specificity 99% [26]. The most frequent sources of error were high myopia mimicking retinitis pigmentosa, which produced false positives, and media opacities, which produced false negatives [26]. A convolutional neural network trained on FAF images reliably distinguished retinitis pigmentosa, Best disease and Stargardt disease [27]. Differentiation of phenocopies is a separate task. A classifier pretrained on FAF images discriminated *ABCA4*-associated Stargardt disease from *PRPH2*-associated pseudo-Stargardt pattern dystrophy, and its performance was compared with that of retinal specialists [28].

Notably, foundation models do not always outperform task-specific architectures. In the classification of IRDs on FAF and ultra-widefield images, ResNet networks achieved a higher weighted F1 score (0.839) than a fine-tuned foundation model [29]. For rare diseases, the choice of architecture therefore remains an empirical question.

3.2. Prediction of genotype from images

The idea of using the phenotype as a genetic test has passed through three stages. The first studies, by a Japanese-British group, classified three genes (*ABCA4*, *RP1L1* and *EYS*) and normal retinas on OCT in a sample of 75 individuals [30]. The same task was then addressed with fundus photographs and FAF images (417 images from 156 individuals), with accuracies of 88.2% and 81.3%, respectively [31]. These studies proved feasibility but did not provide a clinical tool.

The second stage is represented by Eye2Gene. This ensemble of 15 convolutional neural networks was trained on 58 030 FAF, infrared reflectance and OCT images from 2451 patients with confirmed genotypes and predicts 63 genes [32], which together account for more than 90% of IRD cases in European cohorts. In external validation on 775 patients from five centres (Tokyo, Bonn, São Paulo, Oxford and Liverpool), the correct gene was ranked

among the top five in 83% of cases [32]. On FAF images alone, the model identified the correct gene in 76% of cases, whereas experienced clinicians who assessed the same images within the study did so in 36% of cases or fewer. This comparison reflects experimental conditions, in which decisions were based on images without clinical history, and is not equivalent to clinical practice. When used as a phenotypic input for variant prioritisation, the model outperformed tools based solely on phenotype terms in more than 75% of cases [32,33].

The third stage marks the transition to a clinical trial. Retina4IRD is built on RETFound, a foundation model pretrained by self-supervised learning on 1.6 million unlabelled retinal images [34]. The system combines colour fundus photographs, OCT and clinical data, including age at onset, disease duration and family history. It was trained on 1843 patients from nine centres in South Korea, China and Poland and ranks 17 genotype classes selected for the availability of therapies or clinical trials [7,35]. Top-5 accuracy was 90.4% in internal validation and 85.6% in external validation, and discrimination of IRDs from phenotypically similar non-inherited conditions achieved an area under the curve of 0.994 [7].

In a multicentre randomised trial, 300 patients with suspected IRDs were allocated 1:1 to specialists working with or without Retina4IRD support, and sequencing results were obtained for 295 [7]. The primary endpoint was met: top-5 accuracy was 88.5% versus 67.3% ($p < 0.001$), and top-1 accuracy was 37.8% versus 22.4% [7]. The quality of subsequent management also improved, as assessed by decisions on gene therapy, clinical trial enrolment and genetic counselling, and no AI-related adverse events were recorded [7]. This is the first randomised evidence that a decision support system changes clinicians' diagnostic accuracy in a rare inherited disease.

3.3. Quantitative biomarkers and progression

A second line of research concerns the objective quantification of retinal structure. The AIRDetect model automatically segmented the optic disc, areas of hypo- and hyperautofluorescence, the perimacular ring and blood vessels on 45 749 FAF images from 3606 patients with variants in 170 genes [36]. Agreement with expert annotation, measured by the Dice coefficient, ranged from 0.65 for vessels to 0.86 for the optic disc. The model reproduced known genotype-phenotype associations and allowed the rate of ring constriction to be estimated for different retinitis pigmentosa genes [36]. In Stargardt disease, deep learning can annotate several types of FAF lesions simultaneously [37] and increases the sensitivity of fleck detection [38].

In retinitis pigmentosa, the key biomarker is the ellipsoid zone (EZ) area on OCT. Measurements produced by a model and then manually corrected agree with fully manual grading while requiring substantially less time [39], and such models have been tested on baseline natural-history data from *USH2A*-associated retinal degeneration [40]. Prognostic models predict future areas of atrophy in Stargardt disease from OCT and FAF [41]. Using data from 141 eyes in the ProgStar study, recurrent architectures achieved Dice coefficients of approximately 0.88-0.90 for predictions 12-18 months ahead [42]. Finally, models applied to ultra-widefield FAF images from 695 patients estimated visual function, including perimetry and visual acuity [43]. Such biomarkers are needed as endpoints in gene therapy trials and for assessing treatment effects [20,44].

3.4. Technologies addressing data scarcity

The rarity of IRDs results in small and imbalanced datasets, and three technological responses are being explored. The first is foundation models, which are trained without labels and then fine-tuned on small datasets [34]. The second is synthetic data: the SynthEye generative adversarial network produced realistic FAF images for 36 genes, and experts accepted a considerable proportion of the synthetic images as real [45]. A model trained exclusively on synthetic data achieved an area under the receiver operating characteristic curve of 0.86, similar to that of a model trained on real data, but adding synthetic images to real data did not improve classification [45]. The third is federated learning, in which a model is trained across several centres without transferring images [46,47]. Its applicability is limited by data heterogeneity between centres and has so far been demonstrated mainly for common retinal diseases [47,48]. Reviews published in 2025-2026 agree that the technical task is solvable and that the main open questions concern validation and implementation [44,49].

Table 1. Key AI models in inherited retinal dystrophies

Study	Task	Modality	Data	Key result	Level of validation
Fujinami-Yokokawa, 2019 [30]	Gene prediction (3 genes + normal)	OCT	75 individuals, 178 images	High classification accuracy	Internal
Fujinami-Yokokawa, 2021 [31]	Gene prediction (3 genes + normal)	Fundus photography, FAF	156 individuals, 417 images	Accuracy 88.2% (photographs), 81.3% (FAF)	Internal

Miere, 2021 [28]	Stargardt vs pseudo-Stargardt	FAF	370 images	Compared with retinal specialists	Internal
Pontikos (Eye2Gene), 2025 [32]	Ranking of 63 genes	FAF, IR, OCT	2451 patients, 58 030 images	Top-5 83%; FAF alone 76% vs ≤36% for clinicians	External, 5 centres
Jia (Retina4IRD), 2026 [7]	Ranking of 17 genotype classes	Fundus photography, OCT, clinical data	1843 patients; RCT n = 300	Top-5 88.5% vs 67.3% for specialists without AI	RCT, 7 centres
Woof (AIRDetect), 2025 [36]	Feature segmentation	FAF	3606 patients, 45 749 images	Dice 0.65-0.86	Internal, 2 centres
Mishra (ReConNet), 2024 [42]	Prediction of atrophy in Stargardt disease	FAF, OCT	141 eyes (100 patients, ProgStar); OCT, 71 eyes	Dice 0.895 (FAF), 0.882 (OCT)	Internal
Nagasato, 2023 [43]	Estimation of visual function in RP	Ultra-widefield FAF	695 patients, 1274 eyes	Best accuracy with FAF	Internal, 5 centres

Abbreviations: FAF, fundus autofluorescence; IR, infrared reflectance; OCT, optical coherence tomography; RCT, randomised controlled trial; RP, retinitis pigmentosa; top-5, proportion of cases in which the correct gene is among the first five predictions.

4. The translational gap: clinicians and human-AI interaction

4.1. Clinician readiness

No surveys have addressed ophthalmologists' attitudes towards AI specifically in IRDs. The data below therefore concern AI in ophthalmology in general, and extrapolation requires caution. The limitation is important, because specialists in rare diseases and general ophthalmologists may weigh the benefits and risks of such systems differently. An international survey of 622 ophthalmologists from 45 countries, conducted between October 2024 and February 2025, revealed a marked gap between expectations and practice [8]. Although 69.5% of respondents rated the potential of AI to improve outcomes as moderate or high, only 7.2% used it regularly, and 49.6% were unaware of any specific tool (Figure 2). Familiarity with AI was higher among clinicians in academic centres. The main barriers cited were lack of training (20.5%), implementation costs (16.5%) and concerns about reliability (12.9%) [8].

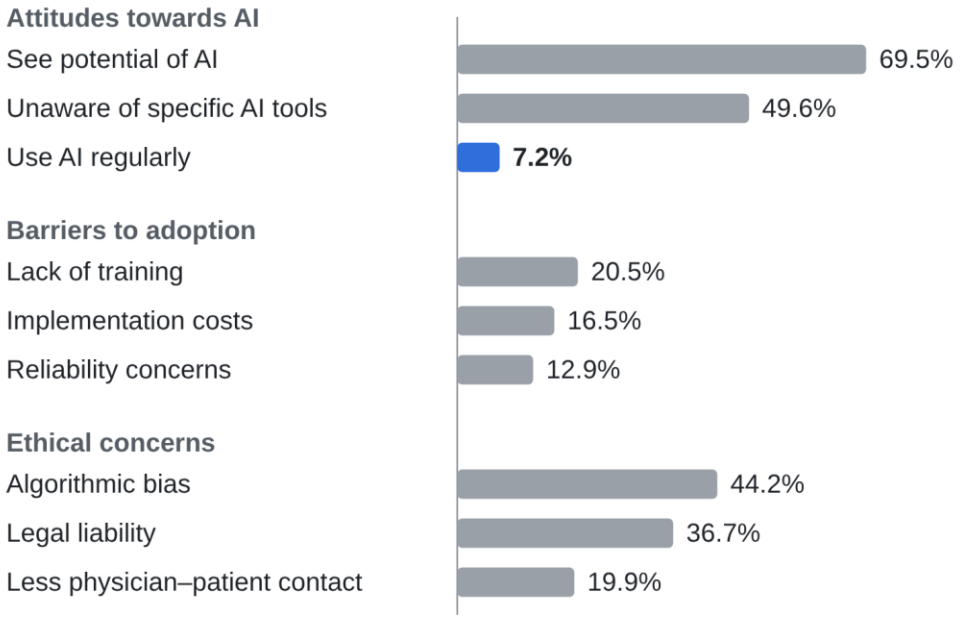

Figure 2. Ophthalmologists' attitudes towards AI, barriers to adoption and ethical concerns: share of respondents (%) in a global survey of 622 ophthalmologists from 45 countries. Data from Ferro Desideri et al., 2025 [8].

What has changed in recent years is not the attitude but the depth of the problem. The APPRAISE survey, conducted in 2020-2021 among 1176 ophthalmologists from 70 countries, found that 88.1% were willing to use AI as an assistive tool [50]. Willingness was higher among clinicians with more than 20 years of experience [50]. A similar effect of experience was confirmed in 2025, when 75.1% of respondents expressed interest in training on AI [8]. The obstacle, therefore, is not rejection of the technology but the lack of accessible tools and skills. This is particularly relevant for IRDs, because the potential users of recognition systems are general ophthalmologists rather than specialists in referral centres.

4.2. Automation bias

In a clinician-AI system, an error made by the algorithm becomes an error made by the clinician. In an experiment with mammography, correct AI suggestions maintained the accuracy of radiologists at approximately 80% [51]. With incorrect suggestions, accuracy fell to 19.8% among inexperienced readers and to 45.5% among the most experienced [51]. Transferring this observation to IRDs requires a caveat about the cost of error. In mammography, an AI error may lead to a missed cancer, whereas in gene ranking it leads to incorrect prioritisation of candidates, which delays diagnosis and may deprive the patient of timely access to a clinical trial. The cognitive mechanism, however, is the same: a clinician who receives a confident but erroneous list of candidates risks anchoring on it when interpreting sequencing results. Even with AI support in the Retina4IRD trial, top-1 accuracy was only 37.8% [7]. The output of such a system should therefore be interpreted as a probabilistic list of hypotheses rather than as a diagnosis.

4.3. Explainability and liability

The ethical concerns most frequently cited by the surveyed ophthalmologists were algorithmic bias (44.2%), legal liability (36.7%) and reduced physician-patient contact (19.9%) [8]. A popular response to mistrust is explainable AI, such as attention heat maps, which Retina4IRD generates [7]. Current explainability methods, however, do not guarantee the correctness of an individual decision and may create a false sense of understanding [52]. Rigorous internal and external validation is a more convincing safeguard than making explainability a precondition for deployment [52]. From a legal standpoint, the least risky role for AI remains that of a confirmatory rather than a determining tool, as long as the standard of care does not explicitly incorporate AI [53]. Finally, issues of privacy and data representativeness are more acute for rare diseases than for common ones [54].

5. Limitations specific to rare diseases

5.1. The closed label set

Any genotype prediction model selects its answer from a finite list. Eye2Gene distinguishes 63 of approximately 280 known genes, and Retina4IRD distinguishes 17 classes defined by therapeutic relevance [7,32]. The architecture of such a classifier does not allow for the answer ‘gene not on the list’, so a patient with a variant in an unrepresented gene will receive a ranking drawn only from the classes known to the model. Whether model confidence falls sufficiently in such cases for them to be recognised has not been systematically assessed in published work, and the developers of Eye2Gene themselves identify limited gene coverage as a limitation of the model [32]. For rare diseases, this is not an incidental flaw but a structural property: the rarest genes are precisely those for which sufficient data will never be available. Synthetic data augmentation did not resolve class imbalance [45]. A clinically responsible system should therefore be able to report that a case lies outside its competence, not merely rank candidates.

5.2. Population shift

Models inherit the structure of the population from which their data were drawn. Eye2Gene was trained predominantly on the Moorfields cohort, which has a European gene spectrum, although it was validated in Tokyo and São Paulo [32]. Retina4IRD was trained on data from South Korea, China and Poland [7]. The leading genes in Japan and the United Kingdom differ fundamentally [9,11], and common population-specific *EYS* variants are almost absent from European databases [12]. The diagnostic yield of sequencing also differs between populations [15]. A systematic review of publicly available ophthalmic imaging datasets showed that most originate from Asia, North America and Europe and that age, sex and ethnicity are poorly reported even at the aggregate level [55]. For IRDs the problem is compounded by the near absence of large molecularly characterised cohorts from Africa [9].

Experience from other fields shows that such shifts are not neutral. Classifiers of chest radiographs systematically underdiagnosed disease in historically underserved groups [56]. When seven commercial diabetic retinopathy screening algorithms were tested on 311 604 real-world images, their sensitivity varied markedly, and one performed worse than human graders in proliferative disease [57]. Metrics obtained on a developer's data do not automatically transfer to another population, another camera or another imaging protocol.

5.3. Limitations of the reference standard

The reference standard for training is a confirmed genetic diagnosis. Yet a quarter to nearly half of patients remain undiagnosed after sequencing [14]. Models trained only on solved cases are biased towards genes that are easier to detect with current methods, while unsolved cases, structural variants and deep intronic variants are under-represented in training data. Moreover, variant interpretation changes over time, and a variant once considered pathogenic may be reclassified [9]. The class label in a training dataset is therefore not absolute.

5.4. Risks of erroneous genomic prediction

A ranking error has concrete consequences for the patient. If a narrow targeted test is ordered on the basis of an AI suggestion instead of a broad panel, the true cause may be missed and the diagnostic odyssey prolonged. An erroneously high probability for a gene with an ongoing clinical trial creates false expectations, whereas a low probability may prematurely close the question of genetic testing. A further risk is vendor dependence: Eye2Gene is integrated into the software platform of a single OCT manufacturer, and models are sensitive to device type and imaging protocol [32]. These considerations define the central principle of the pathway proposed below: AI guides and prioritises but does not narrow genetic testing.

6. An integrated AI-enabled diagnostic pathway

The proposed pathway rests on a single principle: AI shortens the time to the right specialist, enriches the phenotype and increases the yield of genetic diagnosis, but replaces neither sequencing nor clinical judgement (Figure 3). Mass-scale ophthalmology offers two instructive precedents. An autonomous system for detecting diabetic retinopathy was authorised on the basis of a prospective study in 10 primary care practices [58]. In Thailand's national screening programme, an algorithm operating in real time at nine primary care sites achieved an accuracy of 94.7% [59]. Yet the outcome of implementation there was determined not by metrics but by workflows and social context [59], and a pathway for IRDs should be designed with this experience in mind.

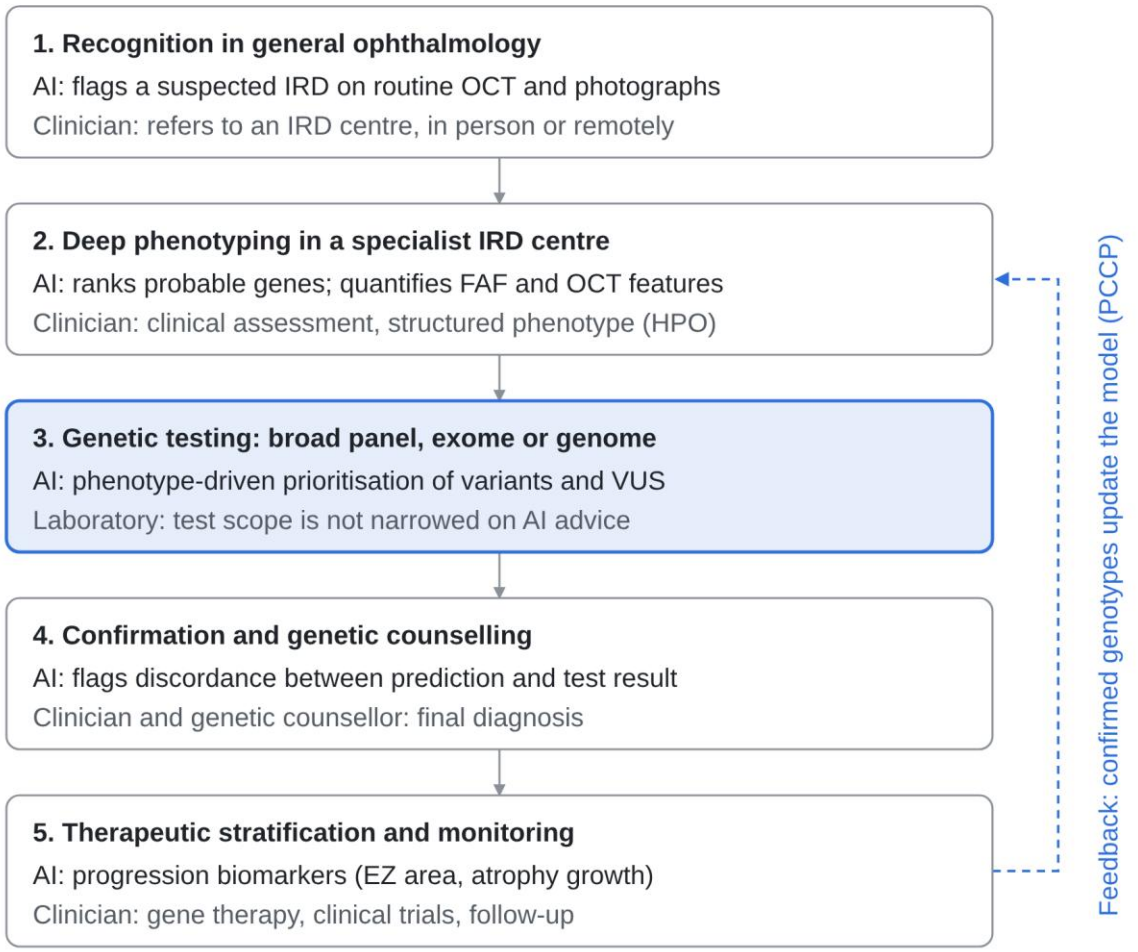

Figure 3. Integrated pathway for a patient with a suspected inherited retinal dystrophy: the role of AI and the decisions retained by clinicians at each stage. The highlighted stage marks the central principle that AI does not narrow the scope of genetic testing; the dashed line shows the feedback loop through which confirmed genotypes update the model under a predetermined change control plan. Authors' scheme. EZ, ellipsoid zone;

FAF, fundus autofluorescence; HPO, Human Phenotype Ontology; IRD, inherited retinal dystrophy; OCT, optical coherence tomography; PCCP, predetermined change control plan; VUS, variants of uncertain significance.

Stage 1. Recognition. This is where AI offers the greatest gain in time. Diagnostic delay is measured in years, and most referrals come from ophthalmologists outside specialist centres [6,17]. A model that distinguishes IRDs from phenocopies (for Retina4IRD, an area under the curve of 0.994 [7]) flags the patient for referral on the basis of routine OCT scans and photographs. Importantly, the system maintained high accuracy with colour fundus photographs alone, the most widely available modality [7].
Stage 2. Deep phenotyping. In the specialist centre, models rank the probable genes and provide a quantitative description of FAF and OCT findings [7,32,36]. The output is recorded as a structured phenotype together with clinical data, which is the form in which the laboratory can use it at the next stage.
Stage 3. Genetic testing. This stage is the key difference between the proposed scheme and the intuitive logic of 'the AI named a gene, so we test only that gene'. A broad panel, exome or genome remains the standard, because its cost is falling while the cost of a missed diagnosis is high (Section 5.4). The role of AI shifts to interpretation. Phenotype-driven variant prioritisation tools such as Exomiser use phenotypes encoded as Human Phenotype Ontology (HPO) terms; in 134 exomes from patients with rare retinal diseases, Exomiser ranked the causative variant among the top five candidates in 94% of cases, compared with 27% without phenotypic data [33]. Exomiser does not analyse images, but image-based AI can serve as its phenotypic input, and in this role Eye2Gene outperformed tools relying on phenotype terms alone [32]. For unsolved cases, discordance between a confident phenotypic prediction and a negative test becomes a reason for targeted reanalysis [16].
Stage 4. Confirmation and counselling. The final diagnosis is established by the clinician together with a genetic counsellor. Discordance between the model prediction and the test result is treated as a reason to revisit the phenotype or the variant classification, not as an automatic decision.
Stage 5. Therapeutic stratification and monitoring. The patient is assessed for approved therapy, for gene-specific trials or, in advanced disease, for mutation-independent approaches [4,20,21]. Automatically measured biomarkers, such as EZ area and the rate of atrophy growth, provide an objective reference point for follow-up and for assessing treatment effects [36,39,42].
Feedback loop. Each confirmed genotype is a new labelled data point. Returning such data to training, including through federated schemes, gradually expands the label set of the model and its population coverage [46,47]. The prerequisite is controlled updating and auditing of performance across subgroups, which is addressed in the next section.

7. System-level enablers

7.1. Regulation of continuously updated models

The feedback loop conflicts with traditional regulatory logic, under which every significant change to medical device software requires a new marketing authorisation. The response has been the predetermined change control plan (PCCP), which the final guidance of the US Food and Drug Administration (FDA), issued in December 2024, extended to all AI-enabled devices [60]. A PCCP comprises a description of the planned modifications, a modification protocol with acceptance criteria and an impact assessment. The guidance also recommends that manufacturers consider the characteristics of the intended population and include measures against bias in data management [60].
In the European Union, the Artificial Intelligence Act, Regulation (EU) 2024/1689, entered into force on 1 August 2024 [61]. AI systems that are, or are part of, medical devices under the Medical Device Regulation or the In Vitro Diagnostic Medical Devices Regulation are automatically classified as high-risk and are subject to dual regulation. The requirements for high-risk AI embedded in such products were originally due to apply from 2 August 2027. The Digital Omnibus package agreed by the Council of the EU and the European Parliament in May 2026 postponed this date to 2 August 2028 [62]. The abstract idea of 'living regulatory sandboxes' has thus acquired concrete instruments. The task for developers of IRD systems is to build population auditing and management of the label set into these instruments.

7.2. Standards of evidence

Existing reporting guidelines define the minimum level of evidence at each stage. CONSORT-AI added 14 items, including descriptions of human-AI interaction and analysis of error cases [63]. DECIDE-AI governs the early clinical evaluation of decision support systems, covering safety, human factors and preparation for larger trials [64]. TRIPOD+AI, a 27-item checklist, specifically requires reporting of fairness across subgroups and of the handling of class imbalance, which is critical for IRDs [65].
The Retina4IRD trial sets the benchmark for the field [7]. General experience, however, warrants caution: in a systematic review of randomised trials of AI-based prediction tools, almost 40% showed no clinical benefit over

standard care [66]. The next generation of studies should assess not only gene-level accuracy but also patient-relevant outcomes, such as time to diagnosis, the proportion of patients enrolled in clinical trials and the timeliness of therapy [67].

7.3. Economics and reimbursement

Who pays for an AI tool is a key practical question and was cited as the second most important barrier [8]. For diabetic retinopathy screening, economic modelling showed that a semi-automated 'AI plus human' model reduced costs by approximately 20% compared with fully manual grading [68], and evidence on the cost-effectiveness of AI in this area continues to accumulate [69]. No comparable evaluations exist for IRDs, which is a substantial gap. Such evaluations are hampered by the rarity of the diseases, the absence of reimbursement codes for AI-based image analysis and the difficulty of placing a monetary value on outcomes such as a shorter time to diagnosis. The sources of savings in IRDs also differ: a shorter diagnostic odyssey, avoidance of repeated uninformative tests and timely enrolment in trials. These should be weighed against a disease burden measured in tens of thousands of US dollars per patient per year [22-24]. A practical way to lower the entry barrier is to integrate models into the software of imaging devices already in use, as has been done for Eye2Gene [32].

7.4. Education

Three-quarters of ophthalmologists are interested in training on AI and prefer online and structured formats [8]. Training programmes should cover not only the capabilities but also the limitations of these systems, including the probabilistic nature of gene ranking, the closed label set and the risk of automation bias [51]. World Health Organization guidance on the ethics and governance of AI in health establishes the principles of human oversight, transparency and accountability, and these principles have been extended to large multimodal models [70,71].

Table 2. Barriers to the implementation of AI in inherited retinal dystrophies and system-level solutions

Barrier	Mechanism	Proposed solution	References
Insufficient training	Clinicians are unaware of available tools and their limitations	AI modules in residency and continuing education; training in recognising model errors	[8,50,51]
Cost and reimbursement	No reimbursement codes; costs of equipment and integration	Economic evaluations for IRDs; integration into imaging platforms; semi-automated models	[32,68,69]
Automation bias	Anchoring on an erroneous suggestion	Output presented as a probabilistic list of hypotheses; clinician remains responsible for the decision	[7,51]
'Black box'	Mistrust of unexplained decisions	Rigorous external validation; heat maps used only as an auxiliary tool	[52]
Population shift	Different gene spectrum, devices and imaging protocols	Multicentre external validation; federated learning; subgroup reporting	[15,46,55,65]
Closed label set	Genes absent from training receive a spurious ranking	Out-of-distribution detection; retention of broad genetic testing	[32,45]
Continuously updated models	Each change requires a new authorisation	PCCP; EU AI Act together with the Medical Device Regulation	[60-62]
Weak outcome evidence	Accuracy assessed instead of patient benefit	CONSORT-AI, DECIDE-AI; pragmatic RCTs with patient-relevant outcomes	[63,64,66,67]
Legal liability	Unclear standard of care	AI as decision support; documentation of discordant findings	[53,70]

Abbreviations: EU, European Union; IRDs, inherited retinal dystrophies; PCCP, predetermined change control plan; RCT, randomised controlled trial.

Conclusion

Within five years, AI in IRDs has progressed from classifying three genes on a few hundred images to a multicentre randomised trial demonstrating that model suggestions improve specialists' diagnostic accuracy. The feasibility of predicting the causative gene from images has been demonstrated, albeit within a limited set of classes and with accuracy far from absolute. The questions that remain open are of a different order: who uses the system and at which stage, how it performs in a population with a different gene spectrum, who is responsible for an error, who pays and how the model can be updated without loss of control.

The main conclusion of this review follows from these questions. The innovation lies not in the algorithm but in its integration into the patient pathway. AI is most valuable in IRDs where it shortens the route to a specialist centre, converts an image into a structured phenotype and increases the yield of genetic diagnosis. It should not be used to narrow genetic testing or as an autonomous source of diagnosis.

Acknowledgments: The authors thank their colleagues for useful discussions and critical comments that helped improve the quality of this review.
Authors' contribution: All authors contributed equally to the conception, design, literature search, analysis and interpretation of the literature, drafting of the manuscript, and critical revision and final approval of the version to be published.
Conflict of interest: The authors declare no conflict of interest.
Financial support: This research received no specific grant from any funding agency in the public, commercial, or non-profit sectors.
Ethics statement: As this article is a narrative literature review and does not involve original data collection from human participants or animals, ethical approval was not required.

References

1. Liew G, Michaelides M, Bunce C. A comparison of the causes of blindness certifications in England and Wales in working age adults (16-64 years), 1999-2000 with 2009-2010. BMJ Open. 2014 Feb 12;4(2):e004015. doi: 10.1136/bmjopen-2013-004015.
2. Hanany M, Rivolta C, Sharon D. Worldwide carrier frequency and genetic prevalence of autosomal recessive inherited retinal diseases. Proc Natl Acad Sci U S A. 2020 Feb 4;117(5):2710-2716. doi: 10.1073/pnas.1913179117.
3. Hanany M, Shalom S, Ben-Yosef T, Sharon D. Comparison of Worldwide Disease Prevalence and Genetic Prevalence of Inherited Retinal Diseases and Variant Interpretation Considerations. Cold Spring Harb Perspect Med. 2024 Feb 1;14(2):a041277. doi: 10.1101/cshperspect.a041277.
4. Russell S, Bennett J, Wellman JA, Chung DC, Yu ZF, Tillman A, Wittes J, Pappas J, Elci O, McCague S, Cross D, Marshall KA, Walshire J, Kehoe TL, Reichert H, Davis M, Raffini L, George LA, Hudson FP, Dingfield L, Zhu X, Haller JA, Sohn EH, Mahajan VB, Pfeifer W, Weckmann M, Johnson C, Gewaily D, Drack A, Stone E, Wachtel K, Simonelli F, Leroy BP, Wright JF, High KA, Maguire AM. Efficacy and safety of voretigene neparvovec (AAV2-hRPE65v2) in patients with RPE65-mediated inherited retinal dystrophy: a randomised, controlled, open-label, phase 3 trial. Lancet. 2017 Aug 26;390(10097):849-860. doi: 10.1016/S0140-6736(17)31868-8. Epub 2017 Jul 14. Erratum in: Lancet. 2017 Aug 26;390(10097):848. doi: 10.1016/S0140-6736(17)32235-3.
5. McClements ME, Elsayed MEAA, Major L, de la Camara CM, MacLaren RE. Gene Therapies in Clinical Development to Treat Retinal Disorders. Mol Diagn Ther. 2024 Sep;28(5):575-591. doi: 10.1007/s40291-024-00722-0.
6. Daich Varela M, Schlottmann P, Luna Pinto J, Michaelides M. Diagnostic Odyssey of More than 1000 Patients with Inherited Retinal Diseases. Ophthalmology. 2024 Feb;131(2):251-253. doi: 10.1016/j.ophtha.2023.10.017.
7. Jia H, Qian B, Qu Y, Wang J, Chen J, Li T, Zheng C, Han J, Zhang G, Jin Y, Lee S, Chen X, Chen H, Xu J, Xu K, Rong Y, Jiang Y, Yang Y, Li W, Guo Q, Li Y, Li Y, Wu J, Zhang Q, Lu F, Zhuang W, Lei B, Byeon SH, Lee J, Smedowski A, Swierczynska M, Zhang D, Rim TH, Grzybowski A, Sheng B, Wong TY, Sun X. AI-based clinician decision support system for diagnosis of inherited retinal diseases: a multicenter, randomized trial. Nat Med. 2026 Sep;32(9):3214-3223. doi: 10.1038/s41591-026-04545-w.
8. Ferro Desideri L, Subhi Y, Roth J, Bernardi E, Melo GB, Tsui E, Fung AT, Sagurski N, Wykoff C, Al-Khersan H, Artiaga JCM, Okonkwo O, Chou HD, Gabrielle PH, Chhablani J, Zinkernagel M, Anguita R. Global perspectives of ophthalmologists on artificial intelligence adoption in clinical practice. Int J Retina Vitreous. 2025 Nov 25;11(1):144. doi: 10.1186/s40942-025-00764-4.
9. Pontikos N, Arno G, Jurkute N, Schiff E, Ba-Abbad R, Malka S, Gimenez A, Georgiou M, Wright G, Armengol M, Knight H, Katz M, Moosajee M, Yu-Wai-Man P, Moore AT, Michaelides M, Webster AR, Mahroo OA. Genetic Basis of Inherited Retinal Disease in a Molecularly Characterized Cohort of More Than 3000 Families from the United Kingdom. Ophthalmology. 2020 Oct;127(10):1384-1394. doi: 10.1016/j.ophtha.2020.04.008.
10. Lin S, Vermeirsch S, Pontikos N, Martin-Gutierrez MP, Daich Varela M, Malka S, Schiff E, Knight H, Wright G, Jurkute N, Simcoe MJ, Yu-Wai-Man P, Moosajee M, Michaelides M, Mahroo OA, Webster AR, Arno G. Spectrum of Genetic Variants in the Most Common Genes Causing Inherited Retinal Disease in a Large Molecularly Characterized United Kingdom Cohort. Ophthalmol Retina. 2024 Jul;8(7):699-709. doi: 10.1016/j.oret.2024.01.012.
11. Suga A, Yoshitake K, Minematsu N, Tsunoda K, Fujinami K, Miyake Y, Kuniyoshi K, Hayashi T, Mizobuchi K, Ueno S, Terasaki H, Kominami T, Nao-I N, Mawatari G, Mizota A, Shinoda K, Kondo M, Kato K, Sekiryu T,

Nakamura M, Kusuhara S, Yamamoto H, Yamamoto S, Mochizuki K, Kondo H, Matsushita I, Kameya S, Fukuchi T, Hatase T, Horiguchi M, Shimada Y, Tanikawa A, Yamamoto S, Miura G, Ito N, Murakami A, Fujimaki T, Hotta Y, Tanaka K, Iwata T. Genetic characterization of 1210 Japanese pedigrees with inherited retinal diseases by whole-exome sequencing. Hum Mutat. 2022 Dec;43(12):2251-2264. doi: 10.1002/humu.24492.
12. Yang L, Fujinami K, Ueno S, Kuniyoshi K, Hayashi T, Kondo M, Mizota A, Naoi N, Shinoda K, Kameya S, Fujinami-Yokokawa Y, Liu X, Arno G, Pontikos N, Kominami T, Terasaki H, Sakuramoto H, Katagiri S, Mizobuchi K, Nakamura N, Mawatari G, Kurihara T, Tsubota K, Miyake Y, Yoshitake K, Iwata T, Tsunoda K; JEGC study group. Genetic Spectrum of EYS-associated Retinal Disease in a Large Japanese Cohort: Identification of Disease-associated Variants with Relatively High Allele Frequency. Sci Rep. 2020 Mar 26;10(1):5497. doi: 10.1038/s41598-020-62119-3.
13. Xiao T, Xie Y, Zhang X, Xu K, Zhang X, Jin ZB, Li Y. Variant Profiling of a Large Cohort of 138 Chinese Families With Autosomal Dominant Retinitis Pigmentosa. Front Cell Dev Biol. 2021 Feb 1;8:629994. doi: 10.3389/fcell.2020.629994.
14. Britten-Jones AC, Gocuk SA, Goh KL, Huq A, Edwards TL, Ayton LN. The Diagnostic Yield of Next Generation Sequencing in Inherited Retinal Diseases: A Systematic Review and Meta-analysis. Am J Ophthalmol. 2023 May;249:57-73. doi: 10.1016/j.ajo.2022.12.027.
15. Lin TY, Wang CY, Chen L, Huang SP. The ethnic disparity in the diagnostic yield of high-throughput next-generation sequencing in inherited retinal diseases: a systematic review and meta-analysis. Ophthalmic Genet. 2025 Jun;46(3):215-224. doi: 10.1080/13816810.2025.2464843.
16. Surl D, Won D, Lee ST, Lee CS, Lee J, Lim HT, Chung SA, Song WK, Kim M, Kim SS, Shin S, Choi JR, Sangermano R, Byeon SH, Bujakowska KM, Han J. Clinician-Driven Reanalysis of Exome Sequencing Data From Patients With Inherited Retinal Diseases. JAMA Netw Open. 2024 May 1;7(5):e2414198. doi: 10.1001/jamanetworkopen.2024.14198.
17. Shah AH, Park E, Luke T, Xu Q, Jewell A, Couser NL. The Role of Genetic Testing in Avoiding Diagnostic Delays in Inherited Retinal Disease. Retin Cases Brief Rep. 2023 Apr 27:10.1097/ICB.0000000000001436. doi: 10.1097/ICB.0000000000001436.
18. Maguire AM, Russell S, Chung DC, Yu ZF, Tillman A, Drack AV, Simonelli F, Leroy BP, Reape KZ, High KA, Bennett J. Durability of Voretigene Neparvovec for Biallelic RPE65-Mediated Inherited Retinal Disease: Phase 3 Results at 3 and 4 Years. Ophthalmology. 2021 Oct;128(10):1460-1468. doi: 10.1016/j.ophtha.2021.03.031.
19. Fischer MD, Simonelli F, Sahni J, Holz FG, Maier R, Fasser C, Suhner A, Stiehl DP, Chen B, Audo I, Leroy BP; PERCEIVE Study Group. Real-World Safety and Effectiveness of Voretigene Neparvovec: Results up to 2 Years from the Prospective, Registry-Based PERCEIVE Study. Biomolecules. 2024 Jan 17;14(1):122. doi: 10.3390/biom14010122.
20. Igoe JM, Lam BL, Gregori NZ. Update on Clinical Trial Endpoints in Gene Therapy Trials for Inherited Retinal Diseases. J Clin Med. 2024 Sep 18;13(18):5512. doi: 10.3390/jcm13185512.
21. Sahel JA, Boulanger-Scemama E, Pagot C, Arleo A, Galluppi F, Martel JN, Esposti SD, Delaux A, de Saint Aubert JB, de Montleau C, Gutman E, Audo I, Duebel J, Picaud S, Dalkara D, Blouin L, Taiel M, Roska B. Partial recovery of visual function in a blind patient after optogenetic therapy. Nat Med. 2021 Jul;27(7):1223-1229. doi: 10.1038/s41591-021-01351-4.
22. Chay J, Tang RWC, Tan TE, Chan CM, Mathur R, Lee BJH, Chan HH, Sim SSKP, Farooqui S, Teo KYC, Fenwick EK, Lamoureux EL, Cheung CMG, Fenner BJ. The economic burden of inherited retinal disease in Singapore: a prevalence-based cost-of-illness study. Eye (Lond). 2023 Dec;37(18):3827-3833. doi: 10.1038/s41433-023-02624-7.
23. Ng QX, Ong C, Yaow CYL, Chan HW, Thumboo J, Wang Y, Koh GCH. Cost-of-illness studies of inherited retinal diseases: a systematic review. Orphanet J Rare Dis. 2024 Feb 29;19(1):93. doi: 10.1186/s13023-024-03099-9.
24. Vandersmissen I, Geers J, Denee T, Oinasmaa S, Hoeben D, Zhang X, Aydin B, Cheung S, Billiet L, Joniau I, Daly A, Simoens S, Leroy BP. The socioeconomic impact of inherited retinal dystrophies (IRDs) in Belgium: A cost-of-illness study. PLoS One. 2026 Jan 27;21(1):e0339332. doi: 10.1371/journal.pone.0339332.
25. Esengönül M, Marta A, Beirão J, Pires IM, Cunha A. A Systematic Review of Artificial Intelligence Applications Used for Inherited Retinal Disease Management. Medicina (Kaunas). 2022 Mar 31;58(4):504. doi: 10.3390/medicina58040504.
26. Ashrafi S, Ahmadieh H, Bayat K, Baghban AA, Azadi A, Shahriari MH, Sabbaghi H. Diagnostic Accuracy of AI Models in Detecting Different Inherited Retinal Diseases: A Systematic Review and Meta-Analysis. Transl Vis Sci Technol. 2025 Jul 1;14(7):22. doi: 10.1167/tvst.14.7.22.
27. Miere A, Le Meur T, Bitton K, Pallone C, Semoun O, Capuano V, Colantuono D, Taibouni K, Chenoune Y, Astroz P, Berlemont S, Petit E, Souied E. Deep Learning-Based Classification of Inherited Retinal Diseases Using Fundus Autofluorescence. J Clin Med. 2020 Oct 14;9(10):3303. doi: 10.3390/jcm9103303.

28. Miere A, Zambrowski O, Kessler A, Mehanna CJ, Pallone C, Seknazi D, Denys P, Amoroso F, Petit E, Souied EH. Deep Learning to Distinguish ABCA4-Related Stargardt Disease from PRPH2-Related Pseudo-Stargardt Pattern Dystrophy. J Clin Med. 2021 Dec 8;10(24):5742. doi: 10.3390/jcm10245742.
29. Trinh H, Ibrahim M, Charng J, Tajbakhsh Z, Chen FK, Mian A, Alam K. Classification of Inherited Retinal Diseases Using Artificial Intelligence Models for Fundus Autofluorescence and Ultrawide Retinal Images. J Ophthalmol. 2026 Jul 31;2026:8810684. doi: 10.1155/joph/8810684.
30. Fujinami-Yokokawa Y, Pontikos N, Yang L, Tsunoda K, Yoshitake K, Iwata T, Miyata H, Fujinami K, Japan Eye Genetics Consortium OBO. Prediction of Causative Genes in Inherited Retinal Disorders from Spectral-Domain Optical Coherence Tomography Utilizing Deep Learning Techniques. J Ophthalmol. 2019 Apr 9;2019:1691064. doi: 10.1155/2019/1691064.
31. Fujinami-Yokokawa Y, Ninomiya H, Liu X, Yang L, Pontikos N, Yoshitake K, Iwata T, Sato Y, Hashimoto T, Tsunoda K, Miyata H, Fujinami K; Japan Eye Genetics Study (JEGC) Group. Prediction of causative genes in inherited retinal disorder from fundus photography and autofluorescence imaging using deep learning techniques. Br J Ophthalmol. 2021 Sep;105(9):1272-1279. doi: 10.1136/bjophthalmol-2020-318544.
32. Pontikos N, Woof WA, Lin S, Ghoshal B, Mendes BS, Veturi A, Nguyen Q, Javanmardi B, Georgiou M, Hustinx A, Ibarra-Arellano MA, Moghul I, Liu Y, Pfau K, Pfau M, Shah M, Yu J, Al-Khuzaei S, Wagner SK, Daich Varela M, Cabral de Guimarães TA, Sen S, Naik G, Sumodhee D, Fu DJ, Kabiri N, Furman J, Liefers B, Lee AY, De Silva SR, Marques C, Motta F, Fujinami-Yokokawa Y, Hardcastle AJ, Arno G, Lorenz B, Herrmann P, Fujinami K, Sallum J, Madhusudhan S, Downes SM, Holz FG, Balaskas K, Webster AR, Mahroo OA, Krawitz PM, Michaelides M. Next-generation phenotyping of inherited retinal diseases from multimodal imaging with Eye2Gene. Nat Mach Intell. 2025;7(6):967-978. doi: 10.1038/s42256-025-01040-8.
33. Cipriani V, Pontikos N, Arno G, Sergouniotis PI, Lenassi E, Thawong P, Danis D, Michaelides M, Webster AR, Moore AT, Robinson PN, Jacobsen JOB, Smedley D. An Improved Phenotype-Driven Tool for Rare Mendelian Variant Prioritization: Benchmarking Exomiser on Real Patient Whole-Exome Data. Genes (Basel). 2020 Apr 23;11(4):460. doi: 10.3390/genes11040460.
34. Zhou Y, Chia MA, Wagner SK, Ayhan MS, Williamson DJ, Struyven RR, Liu T, Xu M, Lozano MG, Woodward-Court P, Kihara Y; UK Biobank Eye & Vision Consortium; Altmann A, Lee AY, Topol EJ, Denniston AK, Alexander DC, Keane PA. A foundation model for generalizable disease detection from retinal images. Nature. 2023 Oct;622(7981):156-163. doi: 10.1038/s41586-023-06555-x.
35. Shanghai General Hospital. An Assistant Model for IRD Care Needs: A Randomized Control Trial. ClinicalTrials.gov identifier: NCT06839170 [Internet]. Bethesda (MD): National Library of Medicine (US); 2025 Feb 21 [updated 2026 Mar 16; cited 2026 Oct 25]. Available from: https://clinicaltrials.gov/study/NCT06839170
36. Woof WA, de Guimarães TAC, Al-Khuzaei S, Daich Varela M, Sen S, Bagga P, Mendes B, Shah M, Burke P, Parry D, Lin S, Naik G, Ghoshal B, Liefers BJ, Fu DJ, Georgiou M, Nguyen Q, Sousa da Silva A, Liu Y, Fujinami-Yokokawa Y, Sumodhee D, Patel P, Furman J, Moghul I, Moosajee M, Sallum J, De Silva SR, Lorenz B, Holz FG, Fujinami K, Webster AR, Mahroo OA, Downes SM, Madhusudhan S, Balaskas K, Michaelides M, Pontikos N. Quantification of Fundus Autofluorescence Features in a Molecularly Characterized Cohort of >3500 Patients with Inherited Retinal Disease from the United Kingdom. Ophthalmol Sci. 2024 Nov 12;5(2):100652. doi: 10.1016/j.xops.2024.100652.
37. Zhao PY, Branham K, Schlegel D, Fahim AT, Jayasundera KT. Automated Segmentation of Autofluorescence Lesions in Stargardt Disease. Ophthalmol Retina. 2022 Nov;6(11):1098-1104. doi: 10.1016/j.oret.2022.05.020.
38. Sabharwal J, Liu TYA, Antonio-Aguirre B, Abousy M, Patel T, Cai CX, Jones CK, Singh MS. Automated identification of fleck lesions in Stargardt disease using deep learning enhances lesion detection sensitivity and enables morphometric analysis of flecks. Br J Ophthalmol. 2024 Aug 22;108(9):1226-1233. doi: 10.1136/bjo-2023-323592.
39. Wang YZ, Juroch K, Birch DG. Deep Learning-Assisted Measurements of Photoreceptor Ellipsoid Zone Area and Outer Segment Volume as Biomarkers for Retinitis Pigmentosa. Bioengineering (Basel). 2023 Dec 6;10(12):1394. doi: 10.3390/bioengineering10121394.
40. Wang YZ, Birch DG. Performance of Deep Learning Models in Automatic Measurement of Ellipsoid Zone Area on Baseline Optical Coherence Tomography (OCT) Images From the Rate of Progression of USH2A-Related Retinal Degeneration (RUSH2A) Study. Front Med (Lausanne). 2022 Jul 5;9:932498. doi: 10.3389/fmed.2022.932498.
41. Mishra Z, Wang Z, Sadda SR, Hu Z. Using Ensemble OCT-Derived Features beyond Intensity Features for Enhanced Stargardt Atrophy Prediction with Deep Learning. Appl Sci (Basel). 2023 Jul 2;13(14):8555. doi: 10.3390/app13148555.
42. Mishra Z, Wang ZC, Xu E, Xu S, Majid I, Sadda SR, Hu ZJ. Recurrent and Concurrent Prediction of Longitudinal Progression of Stargardt Atrophy and Geographic Atrophy towards Comparative Performance on

Optical Coherence Tomography as on Fundus Autofluorescence. Appl Sci (Basel). 2024 Sep 3;14(17):7773. doi: 10.3390/app14177773.
43. Nagasato D, Sogawa T, Tanabe M, Tabuchi H, Numa S, Oishi A, Ohashi Ikeda H, Tsujikawa A, Maeda T, Takahashi M, Ito N, Miura G, Shinohara T, Egawa M, Mitamura Y. Estimation of Visual Function Using Deep Learning From Ultra-Widefield Fundus Images of Eyes With Retinitis Pigmentosa. JAMA Ophthalmol. 2023 Apr 1;141(4):305-313. doi: 10.1001/jamaophthalmol.2022.6393.
44. Chuchmacz W, Bobowska B, Forma A, Dzierżyński E, Puźniak D, Teresińska B, Baj J, Dolar-Szczasny J. The Role of Artificial Intelligence in Imaging-Based Diagnosis of Retinal Dystrophy and Evaluation of Gene Therapy Efficacy. J Pers Med. 2025 Dec 5;15(12):605. doi: 10.3390/jpm15120605.
45. Veturi YA, Woof W, Lazebnik T, Moghul I, Woodward-Court P, Wagner SK, Cabral de Guimarães TA, Daich Varela M, Liefers B, Patel PJ, Beck S, Webster AR, Mahroo O, Keane PA, Michaelides M, Balaskas K, Pontikos N. SynthEye: Investigating the Impact of Synthetic Data on Artificial Intelligence-assisted Gene Diagnosis of Inherited Retinal Disease. Ophthalmol Sci. 2022 Nov 22;3(2):100258. doi: 10.1016/j.xops.2022.100258.
46. Gholami S, Lim JI, Leng T, Ong SSY, Thompson AC, Alam MN. Federated learning for diagnosis of age-related macular degeneration. Front Med (Lausanne). 2023 Oct 12;10:1259017. doi: 10.3389/fmed.2023.1259017.
47. Nabil AS, Gholami S, Leng T, Lim JI, Alam MN. Federated Learning for Multi-Disease Ophthalmic Diagnostics Using OCT Angiography. Ophthalmol Sci. 2025 Dec 18;6(3):101030. doi: 10.1016/j.xops.2025.101030.
48. Wei Y, Zhao K, Grzybowski A, Jin K. Federated learning for privacy-preserving ophthalmic artificial intelligence: clinical applications and translational challenges. Front Med (Lausanne). 2026 Jul 29;13:1904004. doi: 10.3389/fmed.2026.1904004.
49. Kiraly P, Fischer MD. Artificial intelligence applications in inherited retinal dystrophies. Graefes Arch Clin Exp Ophthalmol. 2026 Feb;264(2):299-307. doi: 10.1007/s00417-025-06956-w.
50. Gunasekeran DV, Zheng F, Lim GYS, Chong CCY, Zhang S, Ng WY, Keel S, Xiang Y, Park KH, Park SJ, Chandra A, Wu L, Campbel JP, Lee AY, Keane PA, Denniston A, Lam DSC, Fung AT, Chan PRV, Sadda SR, Loewenstein A, Grzybowski A, Fong KCS, Wu WC, Bachmann LM, Zhang X, Yam JC, Cheung CY, Pongsachareonnont P, Ruamviboonsuk P, Raman R, Sakamoto T, Habash R, Girard M, Milea D, Ang M, Tan GSW, Schmetterer L, Cheng CY, Lamoureux E, Lin H, van Wijngaarden P, Wong TY, Ting DSW. Acceptance and Perception of Artificial Intelligence Usability in Eye Care (APPRAISE) for Ophthalmologists: A Multinational Perspective. Front Med (Lausanne). 2022 Oct 13;9:875242. doi: 10.3389/fmed.2022.875242.
51. Dratsch T, Chen X, Rezazade Mehrizi M, Kloeckner R, Mähringer-Kunz A, Püsken M, Baeßler B, Sauer S, Maintz D, Pinto Dos Santos D. Automation Bias in Mammography: The Impact of Artificial Intelligence BI-RADS Suggestions on Reader Performance. Radiology. 2023 May;307(4):e222176. doi: 10.1148/radiol.222176.
52. Ghassemi M, Oakden-Rayner L, Beam AL. The false hope of current approaches to explainable artificial intelligence in health care. Lancet Digit Health. 2021 Nov;3(11):e745-e750. doi: 10.1016/S2589-7500(21)00208-9.
53. Price WN 2nd, Gerke S, Cohen IG. Potential Liability for Physicians Using Artificial Intelligence. JAMA. 2019 Nov 12;322(18):1765-1766. doi: 10.1001/jama.2019.15064.
54. Lee CS, Brandt JD, Lee AY. Entering the Exciting Era of Artificial Intelligence and Big Data in Ophthalmology. Ophthalmol Sci. 2024 Jan 19;4(2):100469. doi: 10.1016/j.xops.2024.100469.
55. Khan SM, Liu X, Nath S, Korot E, Faes L, Wagner SK, Keane PA, Sebire NJ, Burton MJ, Denniston AK. A global review of publicly available datasets for ophthalmological imaging: barriers to access, usability, and generalisability. Lancet Digit Health. 2021 Jan;3(1):e51-e66. doi: 10.1016/S2589-7500(20)30240-5.
56. Seyyed-Kalantari L, Zhang H, McDermott MBA, Chen IY, Ghassemi M. Underdiagnosis bias of artificial intelligence algorithms applied to chest radiographs in under-served patient populations. Nat Med. 2021 Dec;27(12):2176-2182. doi: 10.1038/s41591-021-01595-0.
57. 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 May;44(5):1168-1175. doi: 10.2337/dc20-1877.
58. Abràmoff MD, Lavin PT, Birch M, Shah N, Folk JC. Pivotal trial of an autonomous AI-based diagnostic system for detection of diabetic retinopathy in primary care offices. NPJ Digit Med. 2018 Aug 28;1:39. doi: 10.1038/s41746-018-0040-6.
59. Ruamviboonsuk P, Tiwari R, Sayres R, Nganthavee V, Hemarat K, Kongprayoon A, Raman R, Levinstein B, Liu Y, Schaekermann M, Lee R, Virmani S, Widner K, Chambers J, Hersch F, Peng L, Webster DR. Real-time diabetic retinopathy screening by deep learning in a multisite national screening programme: a prospective interventional cohort study. Lancet Digit Health. 2022 Apr;4(4):e235-e244. doi: 10.1016/S2589-7500(22)00017-6.

60. U.S. Food and Drug Administration. Marketing Submission Recommendations for a Predetermined Change Control Plan for Artificial Intelligence-Enabled Device Software Functions: Guidance for Industry and Food and Drug Administration Staff [Internet]. Silver Spring (MD): FDA; 2024 Dec 4 [cited 2026 Oct 25]. Available from: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/marketing-submission-recommendations-predetermined-change-control-plan-artificial
61. Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). Official Journal of the European Union. 2024 Jul 12;L 2024/1689 [cited 2026 Oct 25]. Available from: http://data.europa.eu/eli/reg/2024/1689/oj
62. Council of the European Union. Artificial intelligence: Council and Parliament agree to simplify and streamline rules [press release on the Internet]. Brussels: Council of the EU; 2026 May 7 [updated 2026 May 18; cited 2026 Oct 25]. Available from: https://www.consilium.europa.eu/en/press/press-releases/2026/05/07/artificial-intelligence-council-and-parliament-agree-to-simplify-and-streamline-rules/
63. Liu X, Rivera SC, Moher D, Calvert MJ, Denniston AK; SPIRIT-AI and CONSORT-AI Working Group. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI Extension. BMJ. 2020 Sep 9;370:m3164. doi: 10.1136/bmj.m3164.
64. Vasey B, Nagendran M, Campbell B, Clifton DA, Collins GS, Denaxas S, Denniston AK, Faes L, Geerts B, Ibrahim M, Liu X, Mateen BA, Mathur P, McCradden MD, Morgan L, Ordish J, Rogers C, Saria S, Ting DSW, Watkinson P, Weber W, Wheatstone P, McCulloch P; DECIDE-AI expert group. Reporting guideline for the early-stage clinical evaluation of decision support systems driven by artificial intelligence: DECIDE-AI. Nat Med. 2022 May;28(5):924-933. doi: 10.1038/s41591-022-01772-9. Epub 2022 May 18. Erratum in: Nat Med. 2022 Oct;28(10):2218. doi: 10.1038/s41591-022-01951-8.
65. Collins GS, Moons KGM, Dhiman P, Riley RD, Beam AL, Van Calster B, Ghassemi M, Liu X, Reitsma JB, van Smeden M, Boulesteix AL, Camaradou JC, Celi LA, Denaxas S, Denniston AK, Glocker B, Golub RM, Harvey H, Heinze G, Hoffman MM, Kengne AP, Lam E, Lee N, Loder EW, Maier-Hein L, Mateen BA, McCradden MD, Oakden-Rayner L, Ordish J, Parnell R, Rose S, Singh K, Wynants L, Logullo P. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. BMJ. 2024 Apr 16;385:e078378. doi: 10.1136/bmj-2023-078378. Erratum in: BMJ. 2024 Apr 18;385:q902. doi: 10.1136/bmj.q902.
66. Zhou Q, Chen ZH, Cao YH, Peng S. Clinical impact and quality of randomized controlled trials involving interventions evaluating artificial intelligence prediction tools: a systematic review. NPJ Digit Med. 2021 Oct 28;4(1):154. doi: 10.1038/s41746-021-00524-2.
67. Lång K. From algorithms to patient outcomes - lessons from one of the first randomized trials of AI in medicine. Nat Med. 2026 Oct;32(10):3531-3533. doi: 10.1038/s41591-026-04633-x.
68. Xie Y, Nguyen QD, Hamzah H, Lim G, Bellemo V, Gunasekeran DV, Yip MYT, Qi Lee X, Hsu W, Li Lee M, Tan CS, Tym Wong H, Lamoureux EL, Tan GSW, Wong TY, Finkelstein EA, Ting DSW. Artificial intelligence for teleophthalmology-based diabetic retinopathy screening in a national programme: an economic analysis modelling study. Lancet Digit Health. 2020 May;2(5):e240-e249. doi: 10.1016/S2589-7500(20)30060-1.
69. Rajesh AE, Davidson OQ, Lee CS, Lee AY. Artificial Intelligence and Diabetic Retinopathy: AI Framework, Prospective Studies, Head-to-head Validation, and Cost-effectiveness. Diabetes Care. 2023 Oct 1;46(10):1728-1739. doi: 10.2337/dci23-0032.
70. World Health Organization. Ethics and governance of artificial intelligence for health: WHO guidance [Internet]. Geneva: World Health Organization; 2021 [cited 2026 Oct 25]. 150 p. ISBN: 9789240029200. Available from: https://www.who.int/publications/i/item/9789240029200
71. World Health Organization. Ethics and governance of artificial intelligence for health: guidance on large multi-modal models [Internet]. Geneva: World Health Organization; 2024 [cited 2026 Oct 25]. ISBN: 9789240084759. Available from: https://www.who.int/publications/i/item/9789240084759