Received: 5 March 2026 | Accepted: 1 July 2026

DOI: 10.1111/aos.70207

Acta Ophthalmologica

REVIEW ARTICLE

Artificial intelligence prediction algorithms for refractive error onset and progression in children and adolescents: A systematic review and meta-analysis

Athanasia Sandali[1] | **Anna Nikolaidou**[2] | **Theodora Gianni**[3] | **Andreas Katsimpris**[4] | **Ioannis D. Apostolopoulos**[5] | **Lampros Lamprogiannis**[6] | **Eirini Maliagkani**[7]

[1]Postgraduate Program in 'Ocular Surgery', School of Medicine, Aristotle University of Thessaloniki, Thessaloniki, Greece

[2]Institute for Ophthalmic Research, University of Tübingen, Tübingen, Germany

[3]Department of Ophthalmology, 'Aghia Sophia' Children's Hospital, Athens, Greece

[4]Princess Alexandra Eye Pavilion, University of Edinburgh, Edinburgh, UK

[5]Artificial Intelligence, Computational Methods and Technological Applications (ACTA) Lab, University of Thessaly, Larisa, Greece

[6]Ophthalmica Institute of Ophthalmology and Microsurgery, Thessaloniki, Greece

[7]1st Department of Ophthalmology, National and Kapodistrian University of Athens, Athens, Greece

Correspondence
Eirini Maliagkani, 1st Department of Ophthalmology, General Hospital of Athens 'G. Gennimatas', National and Kapodistrian University of Athens, Leoforos Mesogeion 154, 11527 Athens, Greece.
Email: emaliagk@med.uoa.gr; eirini.maliagani@gmail.com

Abstract

This systematic review and meta-analysis evaluates the performance of artificial intelligence (AI)-based models for predicting the onset and progression of refractive error (RE) in children and adolescents and quantitatively synthesizes their prediction accuracy. The study was preregistered in the PROSPERO database (CRD420251075160) and conducted in accordance with PRISMA guidelines. MEDLINE (via PubMed), Web of Science and Scopus were searched without time restrictions. Eligible studies involved children and adolescents (≤18 years) and applied AI-based technologies to predict the onset or progression of RE, using non-image-based input data. To evaluate the prediction accuracy of the included studies, a meta-analysis was performed. Quality assessment was conducted using the PROBAST+AI tool. Sixteen studies were included in the analysis. All studies evaluated myopia, two additionally addressed hyperopia and none focused on astigmatism. Most models predicted continuous SE values across the refractive spectrum, whereas a subset focused on classification-based outcomes, such as myopia onset or progression. Included studies used diverse input data, including age, gender and parental myopia. Eight studies comprising 18 independent evaluations were included in the meta-analysis. The pooled root-mean-squared error (RMSE) for spherical equivalent prediction was 0.50 D (95% confidence intervals: 0.37–0.62). Between-study heterogeneity was substantial (τ^2=0.0695, I^2=100%), indicating major differences across models. This high heterogeneity reflects substantial differences in study populations, prediction targets and horizons, model architectures and validation strategies across included studies, and therefore, the pooled estimate should be interpreted as an average across diverse settings rather than a directly generalizable performance measure. Leave-one-out sensitivity analyses demonstrated that no single study materially influenced the pooled effect. Across all iterations, RMSE estimates ranged narrowly from 0.48 to 0.52 D, with confidence intervals overlapping the main analysis. The studies' quality assessment showed mostly low concern, apart from the analysis domain, in which 43.8% of studies were rated high risk. AI models using non-image-based clinical data demonstrate moderate accuracy for predicting paediatric RE onset and progression. However, substantial methodological heterogeneity and limited external validation indicate that standardized development and validation frameworks are required before clinical implementation.

KEYWORDS
artificial intelligence, children, machine learning, myopia, progression prediction, refractive error

Athanasia Sandali and Anna Nikolaidou contributed equally to this work and shared first authorship.

1 | INTRODUCTION

Refractive errors (REs) such as myopia, hyperopia and astigmatism constitute one of the most prominent causes of paediatric visual impairment worldwide (Cao et al., 2022; Yekta et al., 2022), affecting approximately 12.8 million people aged 5 to 15 years (Resnikoff et al., 2008). REs can often go unnoticed, thus significantly impacting school performance, social development and overall quality of life (Martinez-Perez et al., 2022; Rajabpour et al., 2024).

RE development in childhood is influenced by a complex interplay of biometric, genetic and environmental factors (Harb & Wildsoet, 2019; Zhang & Zou, 2024), particularly in the case of myopia (Martínez-Albert et al., 2023). Established risk factors include axial length (AL), parental myopia and modifiable behaviours, such as near work and limited outdoor activity (Martínez-Albert et al., 2023; Williams et al., 2018). Because myopia often begins early in life and may progress to high myopia with associated sight-threatening complications, timely identification of children at increased risk is critical for effective prevention and management (Williams et al., 2018).

Given the multifactorial nature of RE development, particularly in childhood, artificial intelligence (AI)-based prediction models, including traditional regression approaches as well as machine learning (ML) and neural network–based methods (Choi et al., 2020), provide a framework for integrating diverse demographic, biometric and environmental risk factors. By capturing complex, non-linear relationships and interactions among predictors, and in some cases modelling temporal patterns in longitudinal data, these approaches enable improved prediction of refractive outcomes and facilitate early risk stratification in paediatric populations.

Although several reviews have examined AI applications in RE research, these have largely focused on adult populations or image-based approaches (Ampong et al., 2025; Zhang & Zou, 2024). To our knowledge, no prior study has systematically evaluated and quantitatively synthesized the predictive performance of AI-based models for RE onset and progression exclusively in children and adolescents using non–image-based input data. Therefore, this systematic review and meta-analysis aims to assess the performance of AI-based prediction models for continuous paediatric REs (reported as spherical equivalent [SE]) and to summarize the current evidence regarding their potential clinical utility.

2 | METHODS

2.1 | Study design

This systematic review was registered with the International Prospective Register of Systematic Reviews (PROSPERO; registration number CRD420251075160) and conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Page et al., 2021). The PRISMA checklist is provided in the Supporting Information (Table S1). Given that the study involved the analysis of already published data, ethics approval was not required.

2.2 | Search strategy

A comprehensive search was conducted in the following databases on 14 April 2025: MEDLINE (via PubMed), Web of Science and Scopus. All databases were searched using the following terms: "Artificial Intelligence," "Machine Learning," "Deep Learning," "Neural Network," "Computer Vision," "Refractive Errors," "Myopia," "Hyperopia," "Astigmatism," "Refractive Assessment," "Refractive Progression," "Axial Length," "Child" and "Pediatrics." No time restrictions were applied. The detailed search methodology for each database is provided in the Supporting Information (Table S2). Reference lists of the selected studies were also screened manually. The search was conducted independently by two reviewers (AS, TG).

2.3 | Eligibility criteria & study selection

Eligible studies were peer-reviewed observational (including cohort) or interventional studies (i) involving children and adolescents (≤18 years of age), (ii) using AI-based tools and technologies to predict the onset or progression of REs, (iii) utilizing non-image-based input data, including but not limited to AL, SE, age, sex, parental history, biometric measures, visual acuity, pupil size, screen time or electronic health record data, (iv) reporting model performance metrics (e.g. accuracy, sensitivity and specificity) and (v) published in the following languages, corresponding to the authors' linguistic competencies: English, French, German, Italian, Spanish and Greek. Animal or lab-based studies, studies including only adult populations, non-peer reviewed publications and studies relying exclusively on image-based AI approaches were excluded.

Retrieved records were imported into Rayyan (Ouzzani et al., 2016) for title and abstract screening. Two reviewers (AS and TG), masked to each other, independently screened the titles and abstracts. Duplicate records were removed using Rayyan's deduplication feature prior to screening. No records were excluded by automation tools. Title-abstract and subsequently full-text assessment was performed for eligible studies by the two reviewers (AS and TG), with disagreements resolved by a third reviewer (AN).

2.4 | Data extraction

Data from each eligible study were extracted using a standardized data collection form. Extracted variables included first author, year of publication, country of publication, country of dataset, data source (school, hospital, public or private), number of children and eyes, age range, sex distribution, type of input data (baseline SE, AL, age, sex, parental history, and other clinical or behavioural variables), type of REs (myopia, hyperopia,

etc.), Definitions of REs, use of cycloplegia, reference standard, AI task (classification, regression or progression prediction), AI type, AI algorithm, use of explainable AI (XAI), internal validation method, training, validation and test datasets, external datasets, performance metrics for internal validation and external validation, and limitations as reported by the authors. Where SE was included as an input variable, it referred to baseline RE used as a predictor of future refractive status or progression, rather than the predicted outcome itself. Prediction outcomes were categorized as either continuous (e.g. SE prediction) or classification-based (e.g. myopia onset or progression risk), and this distinction informed the quantitative versus qualitative synthesis approach. Data extraction was performed independently by three reviewers (AS, TG and AN), with final verification conducted by a fourth reviewer (EM).

2.5 | Quality assessment

Quality assessment of the included studies was conducted using the PROBAST+AI tool for prediction models using regression or AI methods (Moons et al., 2025). Each study was assessed independently by two reviewers (AS and TG), with disagreements resolved by a third reviewer (AN). PROBAST+AI evaluates risk of bias across four domains (participants and data sources, predictors, outcomes and analysis) and applicability concerns across three domains (participants, predictors and outcomes) encompassing both model development and model evaluation phases.

2.6 | Statistical analysis

We performed a meta-analysis to evaluate the prediction accuracy of the included studies. All studies included in the meta-analysis, regardless of risk of bias, were retained for quantitative synthesis, and sensitivity analyses were performed to assess the influence of individual studies on the pooled estimates.

The primary effect measure was the root-mean-squared error (RMSE) of SE prediction in dioptres, selected because it is a measure of absolute prediction accuracy and is consistently reported or derivable across most of the included studies. RMSE represents the square root of the mean of squared prediction errors; under the assumption of independent, normally distributed residuals, RMSE has a well-defined sampling distribution that permits construction of standard errors and inverse-variance weighting across studies. When studies reported mean absolute error (MAE) but not RMSE, RMSE was estimated using the relationship $RMSE = MAE \times \sqrt{(\pi/2)}$, which follows from the expected value of the absolute normal distribution and assumes that prediction errors follow a normal distribution with mean zero (Chai & Draxler, 2014; Rice, 2007). By preserving the native dioptre scale and avoiding dependence on prevalence, classification thresholds or scaling transformations, RMSE allowed for meaningful comparison across AI models.

Although the included studies addressed both prediction of continuous RE values and classification-based outcomes (e.g. myopia onset or progression), substantial heterogeneity in outcome definitions and performance metrics precluded quantitative synthesis of classification-based results. In contrast, SE prediction reported as RMSE (or derivable from MAE) was the most consistently available and comparable outcome across studies. Therefore, the meta-analysis was restricted to continuous SE prediction, while classification-based outcomes were summarized qualitatively.

For each study or independent dataset (including distinct external validation cohorts within the same study), a single RMSE estimate was extracted from the held-out test set or external validation dataset. When studies reported performance across multiple independent datasets or validation cohorts, each was included as a separate evaluation in the meta-analysis. The standard error of each RMSE was approximated using the analytic variance of the square root of a chi-square distribution, calculated as $RMSE / \sqrt{(2n)}$, where n is the number of eyes contributing predicted RE measurements.

A random-effects meta-analysis was performed to account for expected methodological heterogeneity among studies. This approach models both within-study sampling variability and between-study heterogeneity in the true prediction error, yielding a pooled RMSE estimate that reflects the average prediction performance across diverse populations and modelling strategies. Between-study heterogeneity was quantified using τ^2 and I^2 statistics. Ninety-five per cent confidence intervals (95% CI) were calculated for the pooled effect to quantify uncertainty around the mean prediction error. In addition, to evaluate the influence of individual studies on the pooled results, we performed a leave-one-out sensitivity analysis.

To explore potential sources of heterogeneity, we conducted subgroup analyses stratified by (i) whether studies used an external dataset for validation, (ii) whether cycloplegic refraction was employed as the reference standard, and (iii) whether RMSE was reported directly or derived from MAE using the conversion formula $RMSE = MAE \times \sqrt{(\pi/2)}$. Differences between subgroups (i) and (ii) were formally tested using meta-regression models. Subgroup analysis (iii) served as a sensitivity analysis to assess whether the MAE-to-RMSE conversion systematically influenced the pooled estimate.

Forest plots were constructed to visualize study-specific RMSE estimates and the pooled effect. All analyses were performed in R (v4.4.0) using the metafor package. A two-sided p-value of less than 0.05 was considered statistically significant.

3 | RESULTS

3.1 | Study selection

The database search identified 530 records, of which 285 remained after duplicate removal. Following title-abstract and full-text screening, 16 studies (Artiemjew et al., 2024; Barraza-Bernal et al., 2023; Huang et al., 2023;

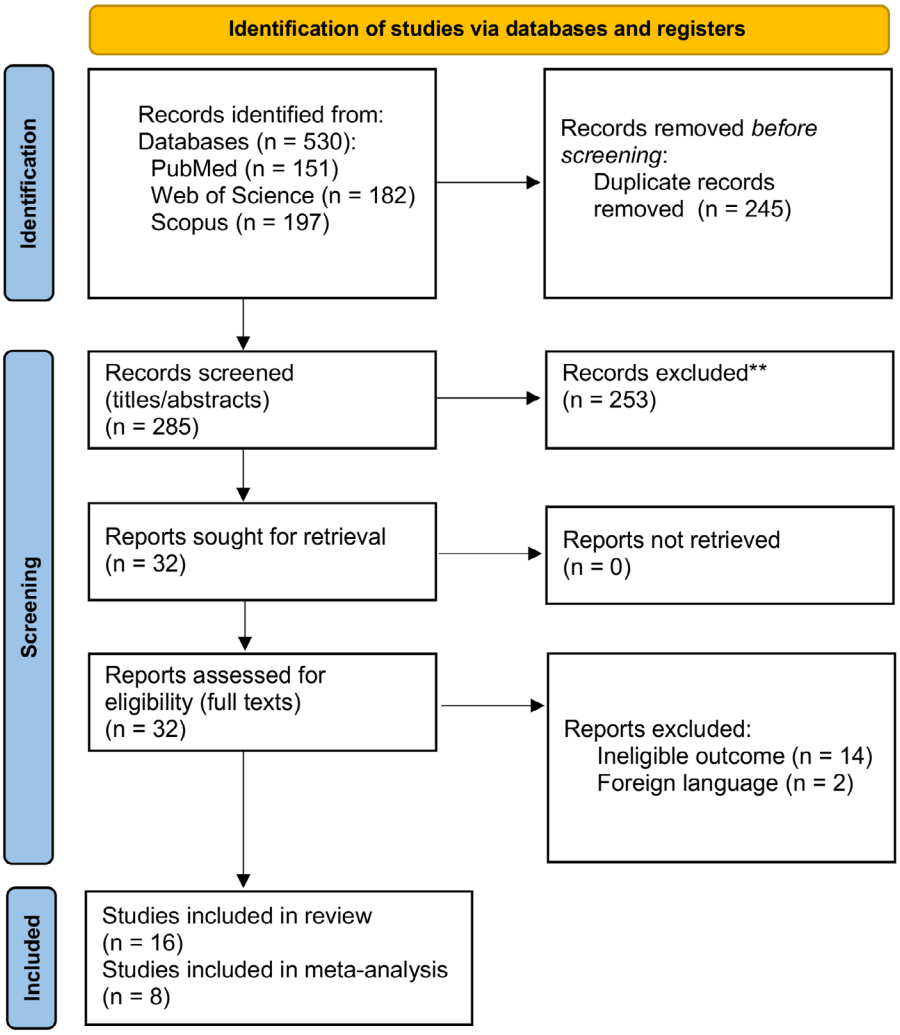

FIGURE 1 Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 Flow Chart, depicting the selection process of the review.

Lai et al., 2024; Li et al., 2022; Li, Tu, et al., 2024; Li, Zeng, et al., 2024; Lin et al., 2018; Mu et al., 2024; Peng et al., 2023; Qi et al., 2024; Varošanec et al., 2024; Wang et al., 2022; Yang et al., 2020; Zhao et al., 2024; Zhu et al., 2023) were selected as per the inclusion criteria. Of these, eight studies (Artiemjew et al., 2024, Barraza-Bernal et al., 2023, Huang et al., 2023, Li, Zeng, et al. 2024; Lin et al., 2018, Varošanec et al., 2024, Zhao et al., 2024, Zhu et al., 2023) were defined as appropriate to include in our meta-analysis. The selection process is depicted in the following flow chart as suggested by PRISMA 2020 (Page et al., 2021) (Figure 1).

3.2 | Study characteristics

A summary of demographic and other fundamental information of the included studies can be found in Table 1. The vast majority of articles primarily studied myopia, with only two studying hyperopia and none studying astigmatism. Included studies utilized a variety of input data to train their algorithms, the most common being age (69%), gender (56%), parental myopia (50%), AL (44%), baseline SE (44%), cycloplegic refraction (31%) and number of myopic parents (31%). In these studies, SE used as an input variable consistently referred to baseline RE, which was incorporated as a predictor of future SE values or RE progression. More information can be found in the Supporting Information (Table S3).

There was also notable heterogeneity concerning the methodology that each study followed when applying AI. As depicted in Table 2, a variety of both ML and deep learning (DL) approaches were followed, including logistic regression, random forest, and even more complicated ensemble models and neural networks. Prediction tasks varied across studies, with most focusing on myopia onset and progression over short or long-term horizons. Explainability techniques were not consistently applied across the included studies, with only seven studies (Li et al., 2022, Li, Tu, et al. 2024; Lin et al., 2018, Mu et al., 2024, Peng et al., 2023, Qi et al., 2024, Zhao et al., 2024) (43.75%) applying SHapley Additive exPlanations (SHAP) or feature-importance plots. Internal validation methods encompassed mainly k-fold cross-validation or fixed training–testing splits, while only five (Li, Tu, et al., 2024; Li, Zeng, et al., 2024; Lin et al., 2018; Peng et al., 2023; Qi et al., 2024) of 16 studies (31.25%) validated their models with external datasets.

TABLE 1 Summary of demographic and other fundamental information of the included studies.

First author (year)	Country	*N*	Age[a]	Sex[b] (F%)	Type of RE	Cycloplegia
Artiemjew et al. (2024)	Poland, Iran, Portugal, USA, Finland	3989	6–12	NM	Myopia	Yes
Barraza-Bernal et al. (2023)	Germany, China	12 780	5–16 (cross-sectional dataset) NM (longitudinal and separated longitudinal datasets)	48.2% (cross-sectional dataset) 40.7% (longitudinal dataset)	Myopia, hyperopia	Yes
Huang et al. (2023)	China	37 586	6–20	48.5%	Myopia	No
Lai et al. (2024)	China, UK	1114	6–8 7.0 ± 0.7	NM	Myopia	No
Li, Zeng, et al. (2024)	China	98 134	8.18 ± 3.54 (training set) 8.14 ± 3.52 (internal validation set)	43.2%	Myopia	Yes
Li et al. (2022)	China, Australia, Belgium, Poland	2740	6–9 7.1 ± 0.4	42.6%	Myopia	Yes
Li, Tu, et al. (2024)	China	15 512	7–10	43.2%	Myopia, emmetropia, hyperopia	Yes
Lin et al. (2018)	China	132 457	8.1 ± 1.5 (training set), 7.4 ± 1.2 (validation set), 6.9 ± 0.5 (multi-source set)	53% (training), 52.4% (validation), 45.9% (multi-source test)	Myopia	Performed using standard protocols in each centre
Mu et al. (2024)	China	2947	6–12	45.1%	Myopia	No
Peng et al. (2023)	China	2221	6–13	48.0%	Myopia	Yes
Qi et al. (2024)	China	1 638 315	4–14 10.10 ± 3.66	47.8%	Myopia	Yes (Cyc-metadata model) No (DeepMyopia model inputs)
Varošanec et al. (2024)	Croatia	895	4–18	58.5%	Myopia	Yes
Wang et al. (2022)	China	15 765	6–18	51.8%	Myopia	No
Yang et al. (2020)	China	3112	6–12	42.2%	Myopia	NM
Zhao et al. (2024)	China	88 250	3–18	48.8%	Myopia	Yes
Zhu et al. (2023)	China	179	6–15	49.7%	Myopia	NM

Abbreviations: *N*, number of children; RE, refractive error.

[a]Age is noted as age range (years) and/or mean age (years ± SD).

[b]Sex is noted as the percentage of female subjects.

TABLE 2 Methodological characteristics of included studies.

First author (year)	AI task	AI type	AI algorithm	XAI	Internal validation method	External dataset
Artiemjew et al. (2024)	Prediction of myopia risk (binary classification) Prediction of spherical equivalent (regression)	ML	Logistic Regression, SVM, k-NN, Gradient Boosting, Artificial Neural Network, Random Forest	No	Monte-Carlo cross-validation; 20 Monte Carlo–simulated runs of 10-fold cross-validation (200 evaluation folds total) combined with bootstrap confidence intervals for metrics	No
Barraza-Bernal et al. (2023)	Estimation of spherical power based on age	ML	Linear Regression, Interaction Linear Regression, Robust Linear Regression, Stepwise Linear Regression, Fine Tree, Medium Tree, Coarse Tree, Linear SVM, Quadratic SVM, Cubic SVM, Fine Gaussian SVM, Medium Gaussian SVM, Coarse Gaussian SVM, Ensemble Boosted Trees, Ensemble Bagged Trees, Squared Exponential GPR, Matern 5/2 GPR, Exponential GPR, Rational Quadratic GPR	No	Models trained on cross-sectional (12 554 children) and longitudinal (226 children, 3 visits) datasets; Hyperparameter optimization performed with a 'separate dataset' to avoid overfitting, but the exact split method/size is not detailed; Final performance evaluated on a hold-out longitudinal set of 81 children not used for training	No
Huang et al. (2023)	Prediction of myopia value in the short and medium term	DL & ML	Time-Aware LSTM, LSTM, Random Forest, Linear Regression	No	80% training, 10% validation, 10% testing	No
Lai et al. (2024)	Prediction of early-onset myopia 2 years after baseline	ML	CART Decision Tree, Naive Bayes, Random Forest, XGBoost, Logistic Regression, Stepwise Logistic Regression, Lasso Logistic Regression	No	Training was done with 10-fold cross-validation; 80% training, 20% testing	No
Li, Zeng, et al. (2024)	Prediction of myopia progression and risk of high myopia	ML	Multivariate Linear Regression, Multivariate Logistic Regression Classifier	No	Training was done with 10-fold cross-validation; 90% training, 10% testing	4 datasets
Li et al. (2022)	Estimation of myopia progression probability and risk-factor analysis	ML	Random Forest	Random Forest Feature Importance	Training was done with 5-fold cross-validation; 90% training, 10% testing	No
Li, Tu, et al. (2024)	Prediction of annual myopia progression and high myopia by risk	ML	XGBoost, k-NN, Decision Tree, Logistic Regression, Gaussian Naive Bayes	XGBoost Feature Importance Plots, SHAP	Training was done with 5-fold cross-validation; 90% training, 10% testing	1 dataset
Lin et al. (2018)	Prediction of future spherical equivalent and onset risk of high myopia at specific future time points	ML	Random Forest	Random Forest Feature Importance, Calibration Plots	10-fold cross-validation & Out-of-bag	9 datasets
Mu et al. (2024)	Classification and prediction of the presence of myopia	ML	Random Forest, Decision Tree, XGBoost, SVM, Logistic Regression	Random Forest Feature Importance	70% training, 30% testing	No
Peng et al. (2023)	Prediction of myopia onset	ML	Random Forest, SVM, Gradient Boosting Decision Tree, CatBoost, Logistic Regression	SHAP	Training was done with 5-fold cross-validation; 70% training, 30% testing	1 dataset

TABLE 2 (Continued)

First author (year)	AI task	AI type	AI algorithm	XAI	Internal validation method	External dataset
Qi et al. (2024)	Prediction of myopia onset, risk stratification and downstream modelling (emulated RCT, lifetime burden)	DL & ML	ResNet-50-based CNN, XGBoost, LightGBM, k-NN, SVM, Random Forest, Artificial Neural Network, Logistic Regression	SHAP (DL model)	70% training, 10% validation, 20% testing	1 dataset
Varošanec et al. (2024)	Prediction of spherical equivalent	DL	Extended Gate Time-Aware LSTM, Time-Aware LSTM, LSTM	No	80% training, 10% validation, 10% testing	No
Wang et al. (2022)	Prediction of myopia and high-myopia risk	ML	Logistic Regression, Random Forest, Gradient Boosted Decision Tree, Artificial Neural Network	No	70% training, 30% testing	No
Yang et al. (2020)	Prediction of myopia status in grade 6 students	ML	SVM, Logistic Regression, Naive Bayes, k-NN, Random Forest, BP Neural Network	No	10-fold cross-validation	No
Zhao et al. (2024)	Prediction of spherical equivalent and onset of myopia and high myopia in a specific year	ML	Random Forest, XGBoost	SHAP	Repeated 5 × 5-fold cross-validation on the training cohort (80% train / 20% validation within each fold) for hyper-parameter tuning	No
Zhu et al. (2023)	Prediction of changes in spherical equivalent refraction and axial length	ML	Orthogonal Matching Pursuit, Random Forest, Kernel Ridge Regression, k-NN Regression, Extra Trees Regression, Multilayer Perceptron	No	70% training, 30% testing	No

Abbreviations: AI, artificial intelligence; CART, classification and regression tree; CNN, convolutional neural network; DL, deep learning; GPR, Gaussian process regression; k-NN, k-nearest neighbours; LSTM, long short-term memory; ML, machine learning; RCT, randomized controlled trial; SHAP, SHapley Additive exPlanations; SVM, support vector machine; XAI, explainable artificial intelligence; XGBoost, extreme gradient boosting.

3.3 | Quality assessment

Methodological quality of the included studies was assessed using the PROBAST+AI tool (Table 3, Figure 2). Across the 16 studies, there was consistently low concern regarding the domains of participants, data sources, predictors and outcomes, with only two studies (Barraza-Bernal et al., 2023; Zhu et al., 2023) showing a high risk of bias in their study design and datasets during model development, one of which (Zhu et al., 2023) also showed a high risk of bias during model evaluation. However, the analysis domain showed significant concerns across the majority of studies, with 43.8% of included studies rated as high risk during both model development and evaluation, primarily due to inadequate handling of missing data, class imbalance or model overfitting. Moreover, an 'unclear' rating in the analysis domain was noted during model development for one quarter of the studies and during model evaluation for half of the studies.

3.4 | Meta-analysis

A total of eight studies were included in the meta-analysis, contributing 18 independent evaluations. The included models were primarily developed to predict continuous RE in SE, rather than exclusively myopia-specific outcomes. Across datasets, the pooled RMSE for SE prediction was 0.50 D (95% CI: 0.37–0.62) (Figure 3), indicating moderate overall prediction accuracy across diverse AI architectures and populations. Between-study heterogeneity was substantial ($\tau^2 = 0.0695$, $I^2 = 100\%$). These differences include variation in baseline RE distributions, age ranges, prediction horizons, geographic location, and whether models were internally or externally validated.

Importantly, despite high statistical heterogeneity, sensitivity analyses demonstrated stable RMSE estimates across leave-one-out iterations, suggesting that the overall findings are robust to the exclusion of individual studies (Supporting Information, Table S4). Across all iterations, RMSE estimates ranged narrowly from 0.48 to 0.52 D, with 95% CIs consistently overlapping the main analysis results. A further sensitivity analysis stratified evaluations by whether RMSE was reported directly or derived from MAE (Supporting Information, Tables S5a–S5c). When restricted to the seven evaluations reporting RMSE directly, the pooled RMSE was 0.29 D (95% CI: 0.14–0.45; $\tau^2 = 0.041$, $I^2 = 100\%$). When restricted to the 11 evaluations in which RMSE was derived from MAE, the pooled RMSE was 0.62 D (95% CI: 0.49–0.75; $\tau^2 = 0.047$, $I^2 = 99.9\%$). Leave-one-out analyses within each subset confirmed stability of the estimates, ranging from 0.23 to 0.32 D in the direct-RMSE subset and from 0.60 to 0.67 D in the converted-RMSE subset. The divergence between subsets reflects clinical variation rather than a systematic conversion artefact: studies reporting RMSE directly reported lower prediction errors (0.13–0.32 D), whereas studies requiring MAE conversion reported higher prediction errors (0.51–0.82 D). Between-study heterogeneity remained substantial in both subsets, consistent with the primary analysis.

Use of an external validation dataset was associated with a higher pooled RMSE (0.57 D vs. 0.35 D), although this difference was not statistically significant ($p = 0.092$) (Supporting Information, Tables S6 and S7). However, models using cycloplegic refraction as the reference standard demonstrated a higher RMSE (0.62 D vs. 0.30 D), and this difference was statistically significant in meta-regression ($p = 0.002$) (Supporting Information, Tables S6 and S7).

4 | DISCUSSION

This systematic review and meta-analysis demonstrates that AI-based models using non–image-derived clinical, biometric and demographic data can predict paediatric RE onset and progression with moderate accuracy. Across 16 studies, the pooled RMSE of approximately 0.50 dioptres suggests that AI methods are capable of generating clinically meaningful predictions. This level of performance was accompanied by substantial methodological heterogeneity and persistent risks of bias in analytical domains. Importantly, sensitivity analyses confirmed that no single study exerted a disproportionate influence on the pooled estimate, supporting the robustness of the findings despite marked heterogeneity. However, the magnitude of heterogeneity limits the interpretability of the pooled estimate and underscores that model performance is highly context-dependent.

An important consideration is that, although the broader objective of this review was to evaluate AI models predicting RE onset and progression, the quantitative synthesis was limited to continuous SE prediction. This reflects the lack of standardized reporting for classification-based outcomes, such as myopia onset or progression risk across studies. As a result, RMSE captures the accuracy of predicting RE as a continuous variable rather than discrete clinical events. While these outcomes are related, they are not equivalent, and this distinction should be considered when interpreting the pooled estimates.

Several studies analysed data at the eye level without accounting for intra-subject correlation from bilaterality and since primary studies did not report the proportion of participants contributing bilateral data or inter-eye correlation in prediction errors, adjustment for within-subject clustering was not feasible at the meta-analytic level. This may have artificially narrowed study-level confidence intervals; however, given that between-study heterogeneity dominated the pooled estimate ($I^2 = 100\%$), the influence on the overall findings is likely limited.

From a clinical perspective, the observed prediction accuracy is unlikely to support refractive prescription; however, it appears well suited for screening and risk stratification. Early identification of children at increased risk of RE development or progression represents a potential application of these models. By integrating routinely collected non-imaging clinical data, AI algorithms may enable clinicians to classify children into low-, moderate- or high-risk categories. Such stratification could facilitate personalized follow-up schedules while reducing unnecessary visits for low-risk individuals.

TABLE 3 Quality assessment of the included studies using the PROBAST+AI tool.

	Quality model development				Risk of bias model evaluation				Applicability model development			Applicability model evaluation		
First author (year)	**D1**	**D2**	**D3**	**D4**	**D1**	**D2**	**D3**	**D4**	**D1**	**D2**	**D3**	**D1**	**D2**	**D3**
Artiemjew et al. (2024)	Low	Low	Low	High	Low	Low	Low	High	Low	Low	Low	Low	Low	Low
Barraza-Bernal et al. (2023)	High	Low	Low	High	Low	Low	Low	High	Low	Low	Low	Low	Low	Low
Huang et al. (2023)	Low	Low	Low	Unclear	Low	Low	Low	High	Low	Low	Low	Low	Low	Low
Lai et al. (2024)	Low	Low	Low	Low	Low	Low	Low	High	Low	Low	Low	Low	Low	Low
Li, Zeng, et al. (2024)	Low	Low	Low	High	Low	Low	Low	High	Low	Low	Low	Low	Low	Low
Li et al. (2022)	Low	Low	Low	High	Low	Low	Low	Unclear	Low	Low	Low	Low	Low	Low
Li, Tu, et al. (2024)	Low	Low	Low	Low	Low	Low	Low	Unclear	Low	Low	Low	Low	Low	Low
Lin et al. (2018)	Low	Low	Low	Low	Low	Low	Low	Unclear	Low	Low	Low	Low	Low	Low
Mu et al. (2024)	Low	Low	Low	Unclear	Low	Low	Low	Unclear	Low	Low	Low	Low	Low	Low
Peng et al. (2023)	Low	Low	Low	Low	Low	Low	Low	Unclear	Low	Low	Low	Low	Low	Low
Qi et al. (2024)	Low	Low	Low	Unclear	Low	Low	Low	Unclear	Low	Low	Low	Low	Low	Low
Varošanec et al. (2024)	Low	Low	Low	Unclear	Low	Low	Low	Unclear	Low	Low	Low	Low	Low	Low
Wang et al. (2022)	Low	Low	Low	High	Low	Low	Low	Unclear	Low	Low	Low	Low	Low	Low
Yang et al. (2020)	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low	Low
Zhao et al. (2024)	Low	Low	Low	High	Low	Low	Low	High	Low	Low	Low	Low	Low	Low
Zhu et al. (2023)	High	Low	Low	High	High	Low	Low	High	Low	Low	Low	Low	Low	Low

Abbreviations: D1, participants and data sources; D2, predictors; D3, outcomes; D4, analyses.

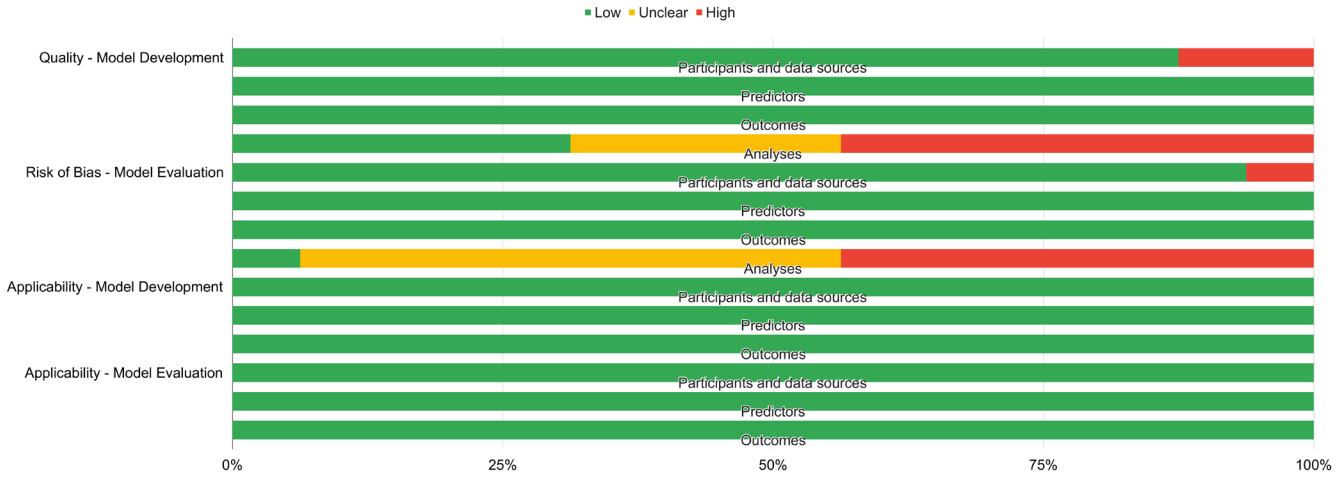

FIGURE 2 Stacked bar chart depicting the quality assessment of the included studies.

In health systems with limited access to paediatric eye care, this targeted approach could help allocate resources more efficiently, reduce the burden of undiagnosed REs and avoid late detection. Children identified as being at elevated risk could benefit from timely lifestyle interventions, including guidance on outdoor activity, near-work habits and screen exposure measures. AI models could also guide early therapeutic interventions aimed at slowing myopia progression, such as low-dose atropine, peripheral defocus spectacle lenses or specialized contact lenses. These treatments have demonstrated efficacy in children with established, progressing myopia (Lawrenson et al., 2025; Maulvi et al., 2025); however, their use in children who have not yet manifested clinically significant RE raises important ethical considerations. Treating an anticipated condition, based on probabilistic model output rather than manifest disease, risks unnecessary exposure to treatment burden.

Compared with prior literature, this review addresses a notable gap by focusing exclusively on paediatric populations and non–image-based AI approaches. Previous reviews and meta-analyses have primarily emphasized adult cohorts (Ampong et al., 2025), limiting their applicability to younger ages, where REs typically emerge and progress. The predominance of myopia-focused studies among the included articles reflects current epidemiologic priorities but also highlights the relative neglect of hyperopia and the complete absence of astigmatism-focused AI prediction models, thereby limiting the comprehensiveness of current evidence. However, global epidemiologic data show substantial and rising prevalence of myopia in children and adolescents worldwide, supporting the notion that myopia is the dominant focus in refractive research (Liang et al., 2025).

Substantial heterogeneity was observed across studies, reflecting wide variation in population characteristics, input features, prediction targets and horizons and AI methodologies. Geographic variation is also relevant, as myopia prevalence and progression differ between East Asian and Western populations. Models developed in high-prevalence settings may not generalize to populations with different risk profiles, contributing to heterogeneity. Although geographic regions may influence model performance, subgroup analysis was not performed due to the small number of non-Asian studies, limiting reliability. Prediction horizons also varied considerably, ranging from short-term forecasts to longer-term progression estimates, which may inherently affect prediction error. Stratified analyses by prediction horizon were not feasible due to inconsistent reporting and limited studies per category. An additional source of heterogeneity relates to differences in prediction targets across studies. While all studies focused on myopia-related research questions, many models were designed to predict continuous RE across the full refractive spectrum, rather than specifically predicting myopia onset or progression as discrete outcomes. This distinction is important, as models predicting SE capture overall refractive change, including emmetropic and hyperopic ranges, whereas classification-based models target clinically defined myopia endpoints. Furthermore, methodological differences, including the use of traditional ML versus DL architectures, as well as differences in validation strategies (internal cross-validation versus independent external validation), represent additional sources of heterogeneity. These factors explain the very high I^2 observed and indicate that the pooled RMSE should be interpreted as a summary measure across heterogeneous contexts rather than a single universally applicable estimate. Furthermore, the sample size of certain models in the included studies varied and in some cases was substantially limited (Barraza-Bernal et al., 2023; Zhu et al., 2023). The sample size may be insufficient, especially considering the datasets are further split for training, validation and testing, thus underlining potential model overfitting and limited generalizability of the results. Differences in model complexity, ranging from traditional ML algorithms to ensemble approaches and neural networks, further contributed to variability in reported performance. Notably, studies that incorporated external validation datasets demonstrated poorer predictive accuracy compared with internally validated models, although this difference did not reach statistical significance.

In addition to methodological variation, several studies reported convergence in the types of input features that contributed most strongly to model performance.

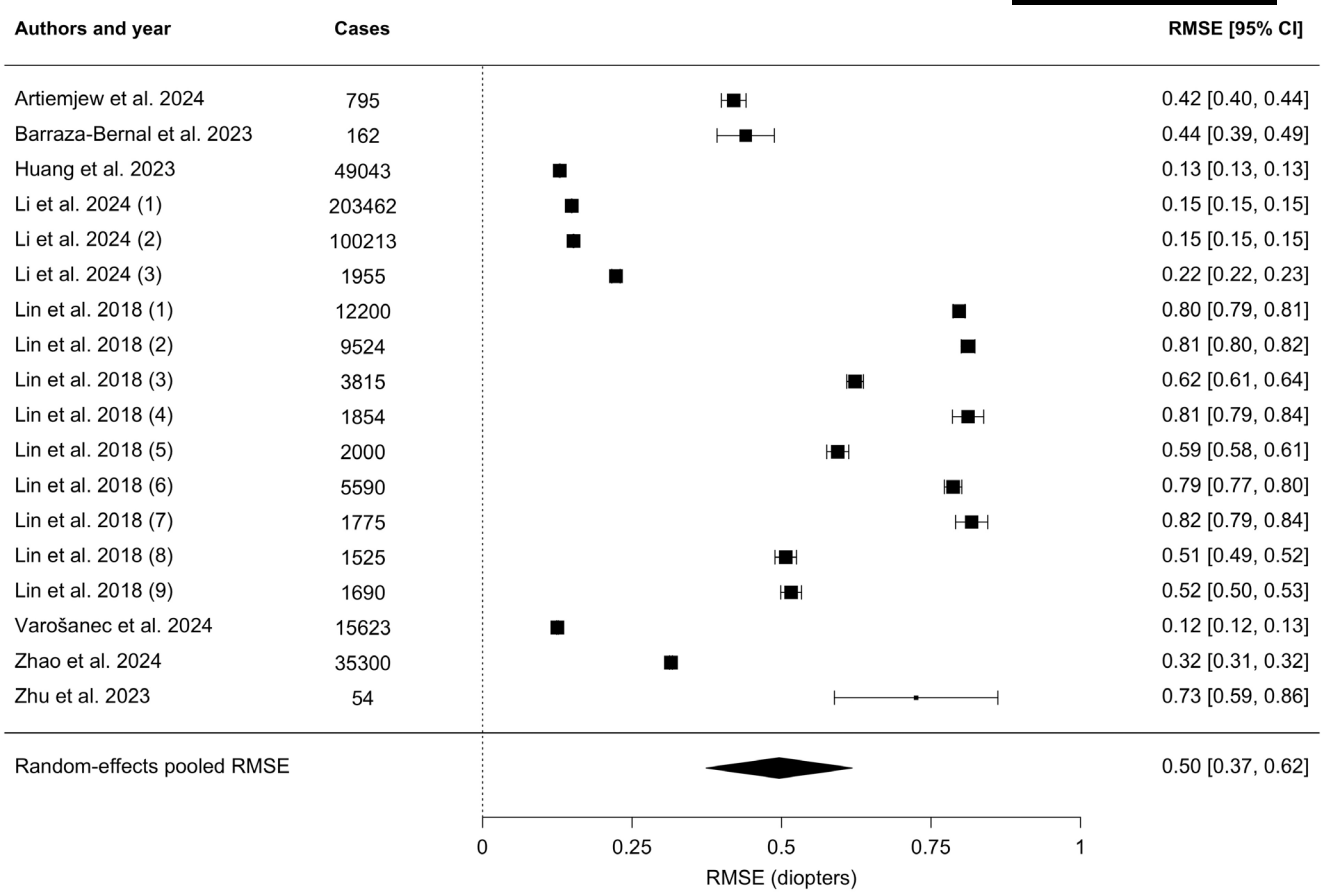

FIGURE 3 Forest plot of study-specific root-mean-squared error (RMSE) values for spherical equivalent (SE) refractive error prediction across 18 model evaluations. Each point represents an individual study or cohort, with horizontal bars indicating 95% confidence intervals (95% CI). Sample sizes (cases) denote the number of eyes included in each evaluation. Two studies contribute multiple entries to the forest plot (Li, Zeng, et al. 2024 [1–3], Lin et al., 2018 [1–9]). These represent independent evaluations derived from distinct datasets within the same study. Each entry was treated as a separate estimate because it reflects model performance in a distinct dataset.

Baseline SE and AL were among the most frequently identified variables, consistent with their established role in refractive development. Demographic and familial factors, including sex and parental myopia, also recurrently contributed to predictive accuracy, while behavioural factors such as reading posture and early onset of myopia were highlighted in multiple models. Collectively, these findings support the biological plausibility of AI-based predictions and suggest that current models predominantly capture well-recognized risk factors rather than identifying novel determinants.

Moreover, shorter prediction horizons were associated with lower prediction error, indicating improved model performance for near-term refractive outcomes. However, although our results suggest that near-term predictions may perform more consistently than longer-range forecasts, the relationship between prediction horizon and error is not clearly established in the existing literature. In Varošanec et al. (Varošanec et al., 2024), variation in prediction accuracy was observed across different historical record lengths and forecast intervals, but no consistent advantage for shorter horizons was demonstrated, indicating that horizon-dependent performance remains uncertain.

An important methodological finding was the significantly poorer performance of models using cycloplegic refraction as the reference standard. Although cycloplegia remains the clinical gold standard for paediatric refractive assessment, its use may introduce greater biological and measurement variability than non-cycloplegic autorefraction, particularly in younger patients with strong accommodative responses (Sankaridurg et al., 2017). As a result, models trained or validated against cycloplegic refraction are effectively calibrated to a potentially noisier outcome measure, which may help explain the higher prediction error observed in these studies. On the other hand, cycloplegia can also be interpreted to reduce accommodative noise, which reflects the true refractive state. Thus, the higher prediction error may be attributed to the model's genuine limitations. This constitutes an important tension, as models evaluated under non-cycloplegic conditions may show better performance due to accommodation-related noise in the reference standard, in contrast to those benchmarked against cycloplegic refraction, which offers higher clinical accuracy assessment.

The overall quality of the included studies was generally acceptable, with low concern in the domains of participants, predictors and outcomes. However, the analysis domain raised important methodological issues. Almost half of the studies were rated as high risk of bias due to limited handling of missing data, class imbalance or potential overfitting. In several studies, key analytical decisions were insufficiently reported, leading to additional 'unclear' ratings. These findings highlight the need for improved transparency and more rigorous

analytical practices in future AI-based prediction studies, particularly with respect to validation procedures and reporting standards. Despite growing interest in AI applications for paediatric RE prediction, explainability remains underutilized. Only a minority of studies employed techniques, such as feature importance analysis or SHAP values. However, established predictors such as baseline SE or AL should not be interpreted as mechanistic discovery, but as a reflection of the model's ability to incorporate known clinical associations. Therefore, it is essential to distinguish biological interpretation from predictive performance and to highlight that accurate predictions do not equal insight into underlying disease mechanisms. Limited model interpretability may hinder clinical trust and adoption, particularly when AI outputs are intended to inform preventive decision-making. This is especially important in paediatric settings, where predictions may be used to guide early interventions before clinically significant refractive deterioration occurs. Improved explainability is therefore an important consideration for safe and responsible integration of AI models into paediatric eye care, as it improves clinician confidence and facilitates transparent communication with families.

In addition to the study-level limitations discussed above, limitations related to this review should also be considered. Although broad database coverage was employed, the review was limited to studies published in English, French, German, Italian, Spanish and Greek, which may have led to the exclusion of relevant evidence published in other languages. Moreover, the extremely high between-study heterogeneity further limits the extent to which pooled estimates can be generalized. Previous research efforts on AI-driven prediction in ophthalmology (Chen et al., 2025; Chen et al., 2026) indicate that between-study heterogeneity varies and is strongly influenced by how limiting the specified inclusion criteria are. This emphasizes the need for standardized reporting and validation practices in future studies. The geographic concentration of available data, with most studies relying on cohorts from East Asian populations, particularly China, constitutes another major limitation which represents an important source of heterogeneity and further constricts the generalizability of the synthesized findings to other regions with differing genetic, environmental and lifestyle profiles (Holden et al., 2016), as patterns of refractive development and myopia progression vary greatly among populations. Additionally, the restriction of the meta-analysis to continuous SE prediction limits direct conclusions regarding the performance of AI models in predicting clinically defined outcomes, such as myopia onset or progression, which are typically evaluated using classification metrics. Moreover, for studies that reported MAE but not RMSE, RMSE was estimated using the transformation $RMSE = MAE \times \sqrt{(\pi/2)}$, which assumes normally distributed residuals with mean zero. If this assumption is violated, the derived RMSE values may not fully reflect the true prediction error, introducing additional uncertainty into the pooled estimates. To address this, a sensitivity analysis stratifying evaluations by conversion method was performed; the divergence between subsets (0.29 D for direct RMSE vs. 0.62 D for converted RMSE) was attributable to clinical variation rather than a systematic bias introduced by the conversion formula, and both subsets demonstrated stable leave-one-out estimates. Finally, as with all evidence syntheses, this review is dependent on the completeness and transparency of reporting in the primary studies, which may have constrained detailed assessment of certain methodological aspects despite the use of a structured quality appraisal tool.

Overall, AI-based prediction models for paediatric REs show promise as tools for early risk stratification rather than diagnosis. Future efforts should prioritize multi-centre, multi-ethnic external validation and transparency to promote models that can be responsibly integrated into paediatric eye care, with careful attention to ethical implications when predictions are used to guide preventive or early interventions in children. Despite these promising applications, an important gap remains between predictive accuracy and real clinical usefulness. Metrics such as RMSE do not show whether model outputs would actually change clinical decisions. Approaches like decision curve analysis could help identify the risk thresholds at which clinicians would adjust follow-up, initiate preventive measures or consider treatment. Moving towards such clinically oriented evaluation will be essential for translating AI models from prediction tools into meaningful decision support systems.

In conclusion, current AI models demonstrate moderate accuracy in predicting paediatric REs using non-image-based clinical data. Heterogeneity in study design, variable methodological quality, lack of external validation and limited model explainability constrain generalizability. Reduced performance observed with cycloplegic refraction and external datasets highlights the importance of standardized inputs, robust validation and data consistency. Overall, these findings support the potential role of AI-based prediction models as adjunct tools in paediatric refractive care, while underscoring the need for careful and responsible integration into clinical practice.

ACKNOWLEDGEMENTS

The publication of this article in OA mode was financially supported by HEAL-Link.

FUNDING INFORMATION

This research did not receive any specific grant from funding agencies in the public, commercial or not-for-profit sectors.

CONFLICT OF INTEREST STATEMENT

The authors have no financial or other conflicts of interest to disclose.

DATA AVAILABILITY STATEMENT

The data extracted and analysed during this systematic review are available from the corresponding author upon reasonable request.

ORCID

Athanasia Sandali https://orcid.org/0009-0002-7706-4654
Andreas Katsimpris https://orcid.org/0000-0002-5805-105X
Ioannis D. Apostolopoulos https://orcid.org/0000-0001-6439-9282
Eirini Maliagkani https://orcid.org/0009-0008-3679-8807

REFERENCES

Ampong, J., Agyekum, S., Eisenbarth, W. et al. (2025) Artificial intelligence applications in refractive error management: a systematic review and meta-analysis. *PLOS Digital Health*, 4, e0000904.

Artiemjew, P., Cybulski, R., Emamian, M.H. et al. (2024) Predicting Children's myopia risk: a Monte Carlo approach to compare the performance of machine learning models. In: *Proc 16th Int Conf agents Artif Intell. Presented at the 16th international conference on agents and artificial intelligence*. Rome, Italy: SCITEPRESS - Science and Technology Publications, pp. 1092–1099.

Barraza-Bernal, M.J., Ohlendorf, A., Diez, P.S., Feng, X., Yang, L.-H., Lu, M.-X. et al. (2023) Prediction of refractive error and its progression: a machine learning-based algorithm. *BMJ Open Ophthalmology*, 8, e001298.

Cao, H., Cao, X., Cao, Z., Zhang, L., Han, Y. & Guo, C. (2022) The prevalence and causes of pediatric uncorrected refractive error: pooled data from population studies for global burden of disease (GBD) sub-regions. *PLoS One*, 17, e0268800.

Chai, T. & Draxler, R.R. (2014) Root mean square error (RMSE) or mean absolute error (MAE)? – arguments against avoiding RMSE in the literature. *Geoscientific Model Development*, 7, 1247–1250.

Chen, K.-Y., Chan, H.-C. & Chan, C.-M. (2025) Can artificial intelligence with multimodal imaging outperform traditional methods in predicting age-related macular degeneration progression? A systematic review and exploratory meta-analysis. *BMC Medical Informatics and Decision Making*, 25, 321.

Chen, K.-Y., Chan, H.-C. & Chan, C.-M. (2026) How do genetic polymorphisms influence the efficacy of age-related macular degeneration treatments? A systematic review and meta-analysis. *Graefe's Archive for Clinical and Experimental Ophthalmology*, 264, 2105–2118.

Choi, R.Y., Coyner, A.S., Kalpathy-Cramer, J., Chiang, M.F. & Campbell, J.P. (2020) Introduction to machine learning, neural networks, and deep learning. *Translational Vision Science & Technology*, 9, 14.

Harb, E.N. & Wildsoet, C.F. (2019) Origins of refractive errors: environmental and genetic factors. *Annual Review of Vision Science*, 5, 47–72.

Holden, B.A., Fricke, T.R., Wilson, D.A. et al. (2016) Global prevalence of myopia and high myopia and temporal trends from 2000 through 2050. *Ophthalmology*, 123, 1036–1042.

Huang, J., Ma, W., Li, R., Zhao, N. & Zhou, T. (2023) Myopia prediction for children and adolescents via time-aware deep learning. *Scientific Reports*, 13, 5430.

Lai, H., Gao, K., Li, M., Li, T., Zhou, X., Zhou, X. et al. (2024) Handling missing data and measurement error for early-onset myopia risk prediction models. *BMC Medical Research Methodology*, 24, 194.

Lawrenson, J.G., Huntjens, B., Virgili, G. et al. (2025) Interventions for myopia control in children: a living systematic review and network meta-analysis. *Cochrane Database of Systematic Reviews*, 2, CD014758.

Li, J., Zeng, S., Li, Z. et al. (2024) Accurate prediction of myopic progression and high myopia by machine learning. *Precis Clinical Medicine*, 7, pbae005.

Li, S.-M., Ren, M.-Y., Gan, J. et al. (2022) Machine learning to determine risk factors for myopia progression in primary school children: the Anyang childhood eye study. *Ophthalmology and Therapy*, 11, 573–585.

Li, W., Tu, Y., Zhou, L., Ma, R., Li, Y., Hu, D. et al. (2024) Study of myopia progression and risk factors in Hubei children aged 7–10 years using machine learning: a longitudinal cohort. *BMC Ophthalmology*, 24, 93.

Liang, J., Pu, Y., Chen, J. et al. (2025) Global prevalence, trend and projection of myopia in children and adolescents from 1990 to 2050: a comprehensive systematic review and meta-analysis. *The British Journal of Ophthalmology*, 109, 362–371.

Lin, H., Long, E., Ding, X. et al. (2018) Prediction of myopia development among Chinese school-aged children using refraction data from electronic medical records: a retrospective, multicentre machine learning study. *PLoS Medicine*, 15, e1002674.

Martínez-Albert, N., Bueno-Gimeno, I. & Gené-Sampedro, A. (2023) Risk factors for myopia: a review. *Journal of Clinical Medicine*, 12, 6062.

Martinez-Perez, C., Alvarez-Peregrina, C., Brito, R. et al. (2022) The evolution and the impact of refractive errors on academic performance: a pilot study of Portuguese school-aged children. *Children*, 9, 840.

Maulvi, F.A., Desai, D.T., Kalaiselvan, P., Shah, D.O. & Willcox, M.D.P. (2025) Current and emerging strategies for myopia control: a narrative review of optical, pharmacological, behavioural, and adjunctive therapies. *Eye*, 39, 2635–2644.

Moons, K.G.M., Damen, J.A.A., Kaul, T. et al. (2025) PROBAST+AI: an updated quality, risk of bias, and applicability assessment tool for prediction models using regression or artificial intelligence methods. *BMJ*, 388, e082505.

Mu, J., Zhong, H. & Jiang, M. (2024) Machine-learning models to predict myopia in children and adolescents. *Frontiers in Medicine*, 11, 1482788.

Ouzzani, M., Hammady, H., Fedorowicz, Z. & Elmagarmid, A. (2016) Rayyan—a web and mobile app for systematic reviews. *Systematic Reviews*, 5, 210.

Page, M.J., McKenzie, J.E., Bossuyt, P.M. et al. (2021) The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*, 372, n71.

Peng, W., Wang, F., Sun, S., Sun, Y., Chen, J. & Wang, M. (2023) Does multidimensional daily information predict the onset of myopia? A 1-year prospective cohort study. *Biomedical Engineering Online*, 22, 45.

Qi, Z., Li, T., Chen, J. et al. (2024) A deep learning system for myopia onset prediction and intervention effectiveness evaluation in children. *npj Digital Medicine*, 7, 206.

Rajabpour, M., Kangari, H., Pesudovs, K., Khorrami-nejad, M., Rahmani, S., Mohaghegh, S. et al. (2024) Refractive error and vision related quality of life. *BMC Ophthalmology*, 24, 83.

Resnikoff, S., Pascolini, D., Mariotti, S.P. & Pokharel, G.P. (2008) Global magnitude of visual impairment caused by uncorrected refractive errors in 2004. *Bulletin of the World Health Organization*, 86, 63–70.

Rice, J.A. (2007) *Mathematical statistics and data analysis*, 3rd edition. Belmont, CA: Duxbury Press.

Sankaridurg, P., He, X., Naduvilath, T. et al. (2017) Comparison of noncycloplegic and cycloplegic autorefraction in categorizing refractive error data in children. *Acta Ophthalmologica*, 95, e633–e640.

Varošanec, A.M., Marković, L. & Sonicki, Z. (2024) A novel time-aware deep learning model predicting myopia in children and adolescents. *Ophthalmology Science*, 4, 100563.

Wang, H., Li, L., Wang, W. et al. (2022) Simulations to assess the performance of multifactor risk scores for predicting myopia prevalence in children and adolescents in China. *Frontiers in Genetics*, 13, 861164.

Williams, K.M., Krapohl, E., Yonova-Doing, E., Hysi, P.G., Plomin, R. & Hammond, C.J. (2018) Early life factors for myopia in the British twins early development study. *The British Journal of Ophthalmology*, 103, 1078–1084.

Yang, X., Chen, G., Qian, Y. et al. (2020) Prediction of myopia in adolescents through machine learning methods. *International Journal of Environmental Research and Public Health*, 17, 463.

Yekta, A., Hooshmand, E., Saatchi, M., Ostadimoghaddam, H., Asharlous, A., Taheri, A. et al. (2022) Global prevalence and causes of visual impairment and blindness in children: a systematic review and meta-analysis. *Journal of Current Ophthalmology*, 34, 1.

Zhang, J. & Zou, H. (2024) Insights into artificial intelligence in myopia management: from a data perspective. *Graefe's Archive for Clinical and Experimental Ophthalmology*, 262, 3–17.

Zhao, J., Yu, Y., Li, Y. et al. (2024) Development and validation of predictive models for myopia onset and progression using extensive 15-year refractive data in children and adolescents. *Journal of Translational Medicine*, 22, 289.

Zhu, S., Zhan, H., Yan, Z., Wu, M., Zheng, B., Xu, S. et al. (2023) Prediction of spherical equivalent refraction and axial length in children based on machine learning. *Indian Journal of Ophthalmology*, 71, 2115.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Sandali, A., Nikolaidou, A., Gianni, T., Katsimpris, A., Apostolopoulos, I.D., Lamprogiannis, L. et al. (2026) Artificial intelligence prediction algorithms for refractive error onset and progression in children and adolescents: A systematic review and meta-analysis. *Acta Ophthalmologica*, 00, 1–14. Available from: https://doi.org/10.1111/aos.70207