



OPEN ACCESS

*CORRESPONDENCE

Guoyuan Yang,
✉ 31981077@qq.com

[†]These authors have contributed equally to this work

RECEIVED 19 July 2024

ACCEPTED 07 November 2024

PUBLISHED 25 November 2024

CITATION

Wu D, Huang X, Chen L, Hou P, Liu L and Yang G (2024) Integrating artificial intelligence in strabismus management: current research landscape and future directions.
Exp. Biol. Med. 249:10320.
doi: 10.3389/ebm.2024.10320

COPYRIGHT

© 2024 Wu, Huang, Chen, Hou, Liu and Yang. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Integrating artificial intelligence in strabismus management: current research landscape and future directions

Dawen Wu^{1,2†}, Xi Huang^{1,3†}, Liang Chen^{1,2}, Peixian Hou^{1,2}, Longqian Liu^{1,2} and Guoyuan Yang^{1,2*}

¹Department of Ophthalmology, West China Hospital, Sichuan University, Chengdu, China,

²Laboratory of Optometry and Vision Sciences, West China Hospital, Sichuan University, Chengdu, China, ³Laboratory of Macular Disease, West China Hospital, Sichuan University, Chengdu, China

Abstract

Advancements in artificial intelligence (AI) are transforming strabismus management through improved screening, diagnosis, and surgical planning. Deep learning has notably enhanced diagnostic accuracy and optimized surgical outcomes. Despite these advancements, challenges such as the underrepresentation of diverse strabismus types and reliance on single-source data remain prevalent. Emphasizing the need for inclusive AI systems, future research should focus on expanding AI capabilities with large model technologies, integrating multimodal data to bridge existing gaps, and developing integrated management platforms to better accommodate diverse patient demographics and clinical scenarios.

KEYWORDS

strabismus, artificial intelligence, deep learning, large model, multimodal, eye care

Impact statement

Early diagnosis and treatment of strabismus are crucial for preventing irreversible visual impairment and improving patient outcomes. This review explores the transformative potential of AI in strabismus management, highlighting significant advancements in screening, diagnosis, and surgical planning that have improved diagnostic accuracy and surgical outcomes. It addresses current challenges such as the underrepresentation of diverse strabismus types and reliance on single-source data, emphasizing the need for inclusive AI systems integrating multimodal data. As the first dedicated review on AI's role in strabismus, it provides valuable insights and guides future research. This new information highlights AI's potential to enhance patient outcomes, improve ophthalmic care, and contribute to societal welfare, setting the stage for further advancements in AI applications in strabismus management.

Introduction

Strabismus, a prevalent ocular disorder characterized by the misalignment of eyes [1], predominantly affects children, leading to rapid deterioration of binocular vision, monocular suppression, anomalous retinal correspondence, and ultimately, irreversible and permanent visual impairment. While primarily seen in children, strabismus can occur at any age, significantly impacting visual function, appearance, learning abilities, employment opportunities [2], and mental health [3], thus constituting a significant societal concern [4]. Delays in diagnosis and treatment often culminate in binocular vision dysfunction or irrevocable vision loss [5]. Given its subtle onset and diverse subtypes, early diagnosis significantly enhances recovery chances. Certain types of infantile strabismus, such as intermittent esotropia in infants under 6 months, may naturally improve by the age of one, advocating for close monitoring to circumvent premature surgical interventions [6]. In light of the above, implementing screening for high-risk groups and precisely diagnosing subtypes of strabismus, coupled with early and appropriate interventions (such as surgery), is of paramount importance. Screening refers to a binary classification process aimed at determining the presence or absence of strabismus, assisting in identifying cases that require referral for further evaluation. Diagnosis, on the other hand, involves multi-class classification to determine the specific subtype of strabismus, which is critical for deciding on appropriate treatment options. However, current strabismus screening and diagnosis are primarily conducted manually by strabismus and pediatric ophthalmologists using methods such as the corneal light reflex test and cover test, which heavily rely on patient cooperation and the physician's skill and experience.

Recent years have seen an exponential growth in artificial intelligence (AI) technology, encompassing research on all common ocular diseases, including anterior segment diseases like keratoconus [7] and cataracts [8], and posterior segment diseases such as retinal diseases [9] and optic nerve-related conditions [10]. The year 2023 witnessed unprecedented breakthroughs in medical AI capabilities, propelled by the transformative development of large models, such as ChatGPT [11]. Google's Med-PaLM [12], achieving expert-level performance on U.S. medical licensing examination questions, and the publication of foundational large model articles in prestigious journals like Nature [13], have underscored the emerging landscape of medical AI large models. The substantial potential of AI has markedly improved the accuracy of ocular disease screening and diagnosis, contributing to reduced healthcare workload, lower medical costs, and addressing the shortage of ophthalmologists, thereby striving for comprehensive healthcare resource coverage and enhanced public health. Notably, AI research in strabismus treatment and prognosis prediction has also been flourishing. This article, set against the backdrop of large model intelligence in medicine,

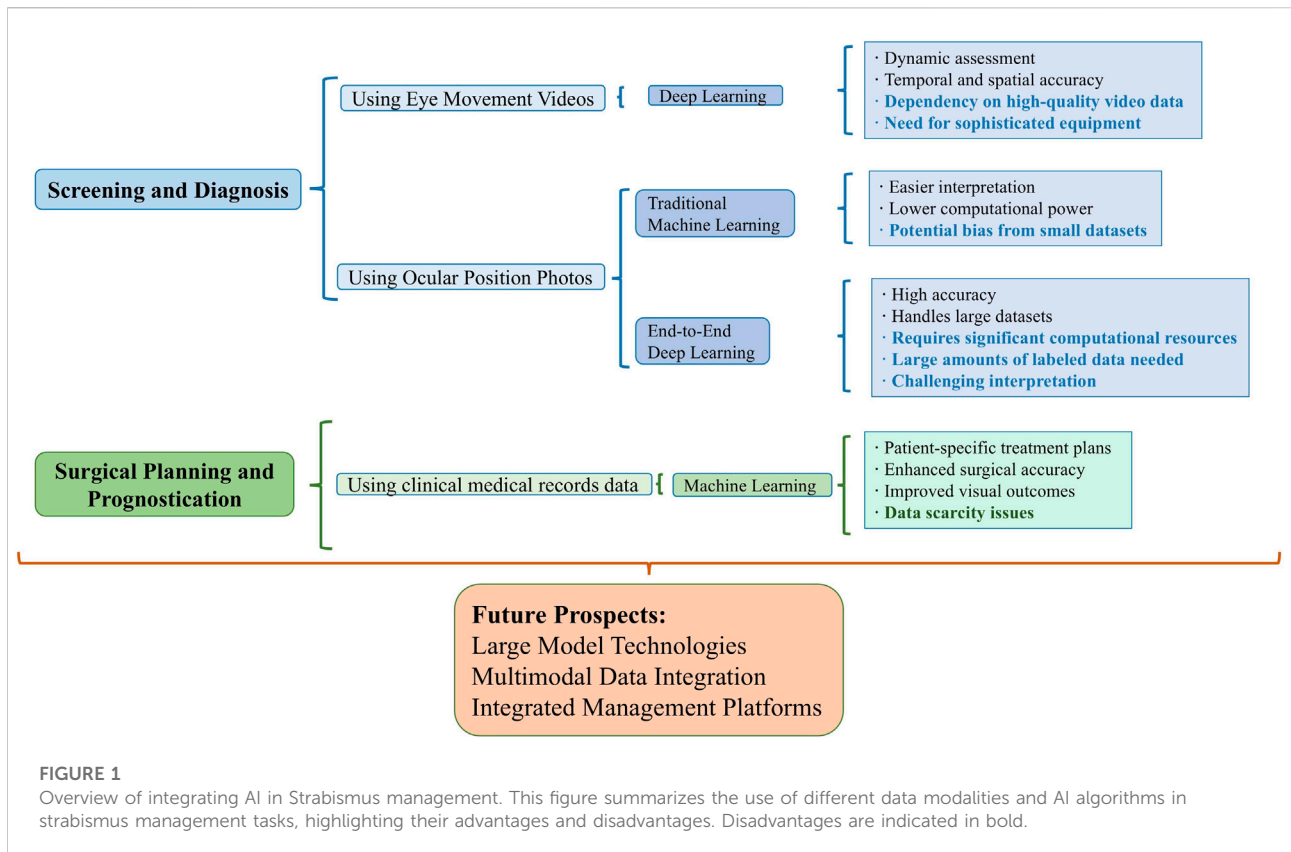
primarily outlines the current state of research on AI applications in the screening, diagnosis, surgical parameter estimation, and prognosis prediction of strabismus, and provides a perspective on its research trends. Figure 1 and Table 1 summarize the relevant studies and applications discussed.

AI in strabismus screening and diagnosis

AI-driven strabismus detection using eye movement videos

The integration of videos with artificial intelligence (AI) in strabismus diagnosis represents a significant advancement in pediatric ophthalmology, offering a nuanced understanding of ocular misalignments through precise and dynamic assessment. This approach capitalizes on the temporal and spatial accuracy of eye tracking to capture subtle deviations in gaze behavior, which are often elusive in traditional examination settings. By harnessing the computational power of AI, these systems analyze complex eye movement patterns efficiently, uncovering diagnostic insights that transcend human observation. Furthermore, the ability to conduct these assessments in a non-invasive, patient-friendly manner reduces the stress associated with conventional diagnostic procedures, facilitating a more comfortable experience for pediatric patients. The amalgamation of eye-tracking technology and AI not only streamlines the diagnostic process but also enhances its accuracy, paving the way for early intervention strategies that are critical in mitigating the long-term visual consequences of strabismus.

In a groundbreaking study published in Nature Biomedical Engineering, Long et al. [14] investigated deep learning for diagnosing visually impaired infants, focusing on strabismus. Using a dataset of over 4,196 infants, they employed a temporal segment network on full-length videos to identify visual impairment patterns. The study highlighted Duane Syndrome and achieved high accuracy with area under the curve (AUC) metrics of 86.4%–93.0% for congenital conditions. This research validates deep learning's potential in diagnosing complex strabismus and underscores video-based behavioral analysis as a non-invasive pediatric ophthalmology tool [14]. Chen et al. [15] published a multicenter study in Nature Medicine, revealing a deep learning-based screening method for pediatric ophthalmic diseases, capable of diagnosing 16 common eye conditions in children, including strabismus. The study used cartoon videos to engage children, while advanced imaging recorded their head and eye movements. The neural network analyzed gaze patterns and facial features, distinguishing specific conditions. Using a dataset of 3,652 subjects and over 25 million video frames, the model achieved an AUC of 0.940 for internal validation and 0.843 for external validation [15].



While the studies mentioned offer significant insights into screening multiple pediatric eye diseases, including strabismus, their accuracy specifically for strabismus diagnosis highlights room for improvement. Given the potential of AI and eye-tracking technologies to advance strabismus diagnosis, automating this process presents a more accurate and efficient alternative to traditional methods. Therefore, previous studies have been published that focus specifically on AI-based strabismus screening and diagnosis using eye movement videos. Miao et al. [16] innovatively apply Virtual Reality (VR) to strabismus diagnosis, introducing a VR-based system that utilizes infrared cameras for measuring ocular deviation. In clinical trials, the VR system achieved a mean deviation of $0.4^\circ \pm 0.2^\circ$ in orthotropic patients and $8.1^\circ \pm 5.5^\circ$ in exotropic patients, closely mirroring doctor evaluations [16]. In the study by Saisara et al. [17], a novel strabismus screening approach combining eye tracking and gaming is proposed through the integration of the Gazepoint GP3 Eye Tracker with custom games. The system was tested on 50 volunteers, effectively identifying strabismus cases with a specific threshold value [17]. Chen et al. [18]'s study employs the Tobii X2-60 eye tracker to automate strabismus diagnosis. By contrasting object positions with eye fixations, their method identifies strabismus types. Precision is ensured through a 25-point calibration test, proving effective in diagnosing conditions like hypertropia [18]. Valente et al. [19] explore

automated strabismus diagnosis through digital video analysis of the cover test. The study outlines an eight-step methodology culminating in a diagnostic accuracy of 93.33% for exotropia, with a specificity of 100% and a sensitivity of 80%. Furthermore, the system reported an average error of 2.57° in deviation measurement [19]. Chen et al. [20] developed a deep learning model capable of diagnosing strabismus by leveraging six different convolutional neural networks (CNNs) - AlexNet, VGG-S, VGG-M, VGG-16, VGG-F, and VGG-19. They collected eye-tracking data as participants looked at nine specific points, represented through gaze deviation (GaDe) images, from 42 subjects for model training and verification. Among the tested networks, VGG-S emerged as the most effective, demonstrating a specificity of 0.960 and sensitivity of 0.941 [20]. These findings highlight the tremendous potential of DL algorithms based on eye movement videos in the field of strabismus screening and subtype diagnosis.

AI-driven strabismus diagnosis using ocular position photos

While eye-tracking technology, with its spatial and temporal accuracy, excels at capturing subtle deviations in gaze behavior and leverages the AI for precise and dynamic evaluation of eye

TABLE 1 Summary of AI applications in strabismus management.

| Authors, year | Task | Age, years | Disease type | Data type | Sample size | AI algorithm | Output |
|---|------------------------------|------------|--|-----------------------------|--------------------------------------|--|--|
| Strabismus Screening and Diagnosis | | | | | | | |
| Long 2019 [14] | Diagnosis | 0–3 | Visual impairment (including strabismus) | videos | 4,196 subjects | Temporal Segment Networks | AUC = 0.816–0.930 |
| Chen 2023 [15] | Screening | 0–4 | Visual impairment (including strabismus) | videos | 3,652 subjects | EfficientNet-B4 | AUC = 0.940 for internal validation AUC = 0.843 for external validation |
| Miao 2020 [16] | Ocular deviation measurement | >6 | Exotropia | VR-based pupil tracking | 17 subjects | stepwise approximation | Mean deviation of $0.4^\circ \pm 0.2^\circ$ in orthotropia, $8.1^\circ \pm 5.5^\circ$ in exotropia |
| Saisara 2017 [17] | Screening | 7–50 | Strabismus | Games and Eye tracking data | 50 subjects | Gazepoint Analysis | Identified Dthreshold ≥ 0.05 for strabismus |
| Chen 2015 [18] | Diagnosis | 3–63 | Esotropia, Exotropia and Hypertropia | Eye tracking data | 225 subjects | Eye-tracking with Tobii X2-60 | Identified subjects' fixation accuracy |
| Valente 2017 [19] | Diagnosis | — | Exotropia | Eye tracking videos | 15 videos of 7 strabismic subjects | Image processing | Sensitivity = 0.800 Specificity = 1.000 Accuracy = 0.933 |
| Chen 2018 [20] | Screening | 25–63 | Recessive, intermittent, and manifest strabismus | Eye tracking images | 42 subjects | AlexNet VGG-F VGG-M VGG-S VGG-16 VGG-17 | VGG-S: Accuracy = 0.952 Specificity = 0.960 Sensitivity = 0.941 |
| Ma 2020 [21] | Screening | 8–10 | Strabismus, myopia and anisometropia | Images | 100 subjects | Image processing | Accuracy = 0.940 Specificity = 0.980 Sensitivity = 0.800 |
| Kang 2022 [22] | Diagnosis | — | Strabismus | Images | 828 subjects | U-Net | Sclera Segmentation: Accuracy = 0.998 Specificity = 0.975 Sensitivity = 0.999 DSC = 0.969 Limbus Segmentation: Accuracy = 0.999 Specificity = 0.956 Sensitivity = 0.999 DSC = 0.957 |
| Almeida 2015 [23] | Diagnosis | — | Exotropia, Esotropia, Hypertropia and Hypotropia | Images | 200 images of 40 strabismic subjects | SVM | Accuracy: 0.880 (ET) 1.000 (XT) 0.803 (HT) 0.833 (HoT) |
| de Oliveira Simoes 2019 [24] | Diagnosis | — | Exotropia, Esotropia, Hypertropia and Hypotropia | Images | 225 images of 45 strabismic subjects | U-Net, ResNet | Accuracy = 0.966 Specificity = 1.000 Sensitivity = 0.958 |
| Mesquita 2021 [25] | Diagnosis | 5–15 | Exotropia, Esotropia, Hypertropia and Hypotropia | Images | 224 subjects | Image processing | Accuracy = 0.845 Specificity = 0.844 Sensitivity = 0.895 Kappa coefficient = 0.430 |
| De Figueiredo 2021 [26] | Diagnosis | 6–87 | Exotropia and Esotropia | Images | | ResNet50 | Accuracy = 0.420–0.920 Precision = 0.250–0.840 |

(Continued on following page)

TABLE 1 (Continued) Summary of AI applications in strabismus management.

| Authors, year | Task | Age, years | Disease type | Data type | Sample size | AI algorithm | Output |
|---|--------------------------|-----------------|---|---------------|---|-------------------------------|--|
| | | | | | 990 images of 110 strabismic subjects | | Recall = 0.380–0.920 F1 = 0.290–0.880 Val_Loss = 0.085–2.210 |
| Huang 2021 [27] | Screening | — | Strabismus | Images | 60 subjects | ResNet-12 | Accuracy = 0.805 Specificity = 0.768 Sensitivity = 0.842 |
| Zheng 2021 [28] | Screening | — | Exotropia and Esotropia | Images | 7,026 images for training and 277 for External Validation | VGG16, Inception-V3, Xception | Inception-V3: AUC = 0.997 Accuracy = 0.968 Specificity = 0.993 Sensitivity = 0.940 |
| Mao 2021 [29] | Screening | 1–65 | Exotropia | Images | 5,797 subjects | InceptionResNetV2 | AUC = 0.998 Accuracy = 0.990 Specificity = 0.983 Sensitivity = 0.991 |
| Wu 2024 [30] | Screening | 1–74 | Exotropia, Esotropia and Vertical deviation | Images | 6,194 images | VIT_16_224 | AUC = 0.994, Accuracy = 0.967, Precision = 0.980, Specificity = 0.970, Sensitivity = 0.960, F1 = 0.975 |
| Kim 2021 [31] | Diagnosis | — | Exotropia and Esotropia | Images | 2023 subjects | CNN | Accuracy = 0.667 |
| Strabismus Surgical Planning and Prognostication | | | | | | | |
| De Almeida 2015 [32] | Surgical planning | — | Exotropia and Esotropia | Clinical data | 88 patients | SVR | Average error: medial rectus muscles: 0.500 mm for recoil, 0.700 for resection lateral rectus muscles: 0.600 for recoil, 0.800 for resection |
| Fernando 2021 [33] | Surgical planning | 0.5–65 | Exotropia and Esotropia | Clinical data | 153 patients | DTR, RFR, ETR | MAE:0.448 – 1.038 mm RMSE = 1.496–2.447 mm |
| Tang 2022 [34] | Surgical planning | — | Strabismus | Clinical data | 1,076 patients | WGAN-GP + lightGBM | AUC = 0.845 |
| Mao 2021 [29] | Surgical planning | 1–65 | Exotropia | Images | 1,070 images | InceptionResNetV2 | accuracy of $\pm 5.5^\circ$ (11.5 PD) with a bias of -0.6° |
| Lou 2023 [35] | Surgical Planning | 17.6 \pm 12.7 | Inferior oblique overaction | Images | 106 eyes | GAR2U-Net | ICC = 0.975 |
| Liu 2019 [36] | Surgical Prognostication | 6–12.25 | Intermittent exotropia | Clinical data | 132 patients | SVM | Accuracy = 0.821 |

AI, artificial intelligence; AUC, area under curve; DSC, dice similarity coefficient; SVM, support vector machine; ET, esotropia; XT, exotropia; HT, hypertropia; HoT, hypotropia; CNN, convolutional neural network; SVR, support vector regression; DTR, decision tree regressor; RFR, random forest regressor; ETR, extra trees regressor; MAE, mean absolute error; RMSE, root mean squared error; LightGBM, light gradient-boosting machine; PD, prism diopter; ICC, intraclass correlation coefficient. This table summarizes various studies on AI applications in strabismus management, detailing the author and year, task, data type, sample size, AI algorithm, and key performance metrics.

movements in both “resting” and “task” states, facilitating the screening of multiple ocular diseases and identification of certain strabismus subtypes, the collection of eye movement videos often requires specialized equipment and is more time-consuming. Additionally, video processing demands substantial computational resources, unlike the convenience and lower computational requirements of ocular position photos. Merely

a camera or even a smartphone can complete the image capture process. Instant uploads can return diagnostic results within milliseconds. Therefore, specifically for strabismus, AI systems based on ocular position photos continue to play a unique and indispensable role.

In the current realm of AI research for strabismus screening and diagnosis using ocular position photos, two main approaches

prevail: algorithms for key eye region segmentation based on traditional stepwise learning, and classification algorithms based on end-to-end learning.

Traditional machine learning algorithms

Ma et al. [21] developed a smartphone app for the rapid screening of pediatric eye diseases, including strabismus. Utilizing image processing and AI, the app analyzes photos taken in dark rooms to evaluate facial landmarks, head tilt, and eye positions. It employs shape fitting techniques to estimate corneal light reflex and red reflex contours, quickly assessing risks of strabismus, myopia, and refractive errors in just 10 s. The app's sensitivity and specificity for detecting strabismus are 0.80 and 0.98, respectively, proving effective for early diagnosis [21]. Kang et al. [22] developed a deep learning model using U-Net architecture to detect strabismus by segmenting the cornea and scleral limbus from 828 gaze photographs across nine different gaze positions. The model demonstrated high segmentation accuracy: 0.9984 for the cornea and 0.9992 for the limbus, with Dice Similarity Coefficients of 0.9688 and 0.9571, respectively [22]. Almeida et al. [23] employed support vector machines (SVMs) for a machine learning-based diagnostic approach to strabismus. Analyzing 200 images from 40 patients with conditions like esotropias, exotropias, hypertropias, and hypotropias, their methodology involved face segmentation and eye region detection to precisely identify eye deviations. It achieved accuracy rates of 88% for esotropias, 100% for exotropias, and over 80% for vertical deviations, with errors closely matching specialist assessments [23]. de Oliveira Simoes et al. [24] adopted a similar experimental approach. Employing a private dataset of 225 images from 45 patients, the research aimed at evaluating ocular alignment by analyzing the distance between the limb centroid and a point between the eye corners, achieving an impressive accuracy of 96.6%, with a sensitivity of 95.8% and specificity of 100% [24]. Mesquita et al. [25] explored the concordance between expert ophthalmologist diagnoses and those made by an mHealth application for strabismus in 224 children aged 5–15. The app analyzed smartphone photos, incorporating steps like face segmentation, eye region detection, and alignment comparison of the Limbus center with brightness location. The app demonstrated a sensitivity of 89.47% and specificity of 84.39% at 6 PD, affirming its utility as a screening tool despite some misclassifications [25]. De Figueiredo et al. [26] developed an app based on the ResNet50 neural network to diagnose strabismus by identifying patients' different gaze positions in photographs from 110 patients. The model, trained to recognize combinations of left and right eyes in gaze positions ranging from 1 to 9 and version classifications from -4 to +4, achieved an overall accuracy between 0.42 and 0.92 and precision between 0.28 and 0.84 [26]. Huang et al. [27] developed a strabismus screening model with frontal facial images for face and eye region identification, and refining the detection with Otsu's binarization

and the HSV color model. The model uses the least squares method to locate the pupil's center and assesses strabismus through eye positional similarity. Additionally, utilizing the ResNet-12 network for screening on images from 60 subjects, the team achieved diagnostic accuracy, sensitivity, and specificity of 0.805, 0.768, and 0.842, respectively [27].

End-to-end deep learning algorithms

Zheng et al. [28] leveraged the Inception-v3 architecture to train a model for identifying referable horizontal strabismus in children's primary gaze photos, excluding conditions like vertical and paralytic strabismus. It achieved an average AUC of about 0.99, with sensitivity at 94.0%, specificity at 99.3%, and accuracy at 96.8% in the external validation set, outperforming resident ophthalmologists [28]. Lin et al. [29] utilized the InceptionResNetV2 architecture for training a model aimed at diagnosing horizontal strabismus, deliberately excluding vertical strabismus and intermittent exotropia during the data collection phase. The training involved 1,561 photographs of horizontal strabismus (both esotropia and exotropia) and 2,496 of normal eyes. Tested on 356 horizontal strabismus and 514 normal eyes, the model achieved a sensitivity of 0.991, a specificity of 0.983, and an accuracy of 0.990 [29]. Wu et al. [30] constructed the largest corneal light-reflection photo dataset in the field to date and trained a model based on the Transformer architecture (ViT_16_224). Unlike previous studies that focused solely on pediatric patients and a limited number of strabismus subtypes, this study encompassed all age groups and a comprehensive range of strabismus subtypes, resulting in the development of the best-performing model. The ViT_16_224 architecture outperformed the models from the aforementioned studies on a shared dataset. On an independent test set, the model achieved an accuracy of 0.967, precision of 0.980, specificity of 0.960, sensitivity of 0.970, and an F1 score of 0.975, significantly improving the generalizability and practicality of the diagnostic model. Donghwan Kim et al. [31] introduced a Convolutional Neural Network (CNN)-based model for classifying strabismus into three categories: esotropia, exotropia, and normal eye alignment, leveraging a "9-photo" front view. The training set consisted of "9-photo" sets from 73 esotropia patients, 75 exotropia patients, and 72 individuals with normal alignment. Testing was conducted with "9-photo" sets from 10 patients in each group. The model achieved a final test accuracy of 66.7% [31]. These studies highlight the model's potential to streamline early strabismus diagnosis in young children.

Comparison of traditional machine learning and end-to-end deep learning algorithms

In comparing traditional machine learning algorithms and end-to-end deep learning algorithms for strabismus diagnosis, several key differences and similarities emerge. Traditional machine learning algorithms typically involve a series of pre-

processing steps, feature extraction, and model training phases. Common techniques include Support Vector Machines (SVM), decision trees, and regression models. For eye region segmentation, researchers often use pre-trained facial detection models to extract key regions and calculate coordinates for critical areas like the pupil center and corneal light reflex points. By comparing these results to predetermined thresholds, the presence and subtype of strabismus can be determined. These methods are easier to interpret and require less computational power, allowing for fine-tuning at each stage. However, their performance heavily depends on the quality and relevance of manually extracted features and may struggle with high-dimensional data and complex patterns inherent in ocular position photos. Additionally, the selection of thresholds based on limited statistical data within smaller datasets may introduce bias, reducing the model's generalizability.

In contrast, end-to-end deep learning algorithms, such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs), learn features and patterns directly from raw data, automating the feature extraction process. These models use extensive image datasets to enhance the model's generalization capability, handling large, high-dimensional datasets and capturing complex, non-linear relationships within the data. This leads to higher accuracy and robustness in predictions. However, they require significant computational resources and large amounts of labeled data for training, and their "black-box" nature can make interpretation and troubleshooting more challenging.

In conclusion, end-to-end deep learning models generally outperform traditional machine learning algorithms in accuracy and generalizability due to their ability to learn directly from data without manual feature extraction. Deep learning models scale better with increased data, improving performance as more data becomes available, whereas traditional methods may not show the same level of improvement due to the limitations of manual feature extraction. While traditional machine learning models are more interpretable, providing clearer insights into the decision-making process, deep learning models offer superior performance and scalability for strabismus diagnosis using ocular position photos. The choice between the two depends on the specific requirements and constraints of the clinical application, including the availability of computational resources and the need for model interpretability. This comparative understanding highlights the potential of integrating both approaches to optimize diagnostic accuracy and clinical applicability.

AI in strabismus surgical planning and prognostication

The application of AI in strabismus surgical planning and prognostication represents a transformative shift towards precision medicine in pediatric ophthalmology. Recent studies

exemplify the integration of advanced machine learning models, such as Support Vector Regression (SVR) [37], multi-output regression trees [38], Light Gradient Boosting Machines (LightGBM) [34], convolutional neural networks (CNN) [39], and recurrent residual CNNs [40] with global attention gates, to refine surgical strategies for complex ocular deviations. These AI-driven tools process extensive clinical data, including visual acuity, type of deviation, binocular fixation, and ocular position, to predict the precise amounts of muscle resection and recoil needed in surgeries.

De Almeida et al. [32] conducted a regression study on 88 patients with horizontal strabismus using Support Vector Regression (SVR). The model aimed to accurately estimate the adjustments needed for medial and lateral rectus muscles. The SVR model achieved mean absolute errors of 0.5 mm for medial rectus recoil and 0.7 mm for resection, with 0.6 mm for lateral rectus recoil and 0.8 mm for resection [32]. Leite et al. [33] analyzed surgical planning for strabismus in 153 patients with horizontal deviations, primarily esotropia and exotropia. Utilizing comprehensive patient data including age, binocular fixation, and deviation from five ductions positions, the study applied a multi-output regression tree approach to predict necessary surgical adjustments for various muscles. The model accurately forecasted recoil and resection requirements for both medial and lateral rectus muscles, achieving a Mean Absolute Error (MAE) range of 0.448 mm–1.038 mm and a Root Mean Square Error (RMSE) of up to 1.496 mm [33]. Tang et al. [34] conducted a study involving 1,076 surgical cases of patients with various types of strabismus, including esotropia and exotropia, using patient data like age, dominant eye, and exodeviation angles. They developed a model using the Light Gradient Boosting Machine (LightGBM) architecture, which initially achieved a 69.32% accuracy in predicting surgical adjustments such as muscle movement amount and direction. To address data scarcity and improve model performance, the Wasserstein Generative Adversarial Network with Gradient Penalty (WGAN-GP) was applied, enhancing accuracy to 84.52% [34]. Mao et al. [29] developed an Operation Advice System using corneal light-reflection photos from 56 exotropia patients who had successfully undergone initial surgeries. The retrospective test set included 160 subjects, predominantly with intermittent exotropia, and demonstrated a high positive correlation ($r = 0.86$, $P < 0.001$) between the predicted and actual surgical angles, with an accuracy of $\pm 5.5^\circ$ (11.5 PD) and a slight bias of -0.6° [29]. Lou et al. [35] developed a recurrent residual CNN based on GAR2U-Net to evaluate Inferior Oblique Overaction (IOOA) in a study involving 106 eyes of 72 patients. The technique focused on measuring the height difference between the inferior corneal limbus of both eyes, a key indicator of IOOA, which was clinically graded from +1 to +4. The study found significant correlations between automated photographic measurements and clinical gradings, and excellent agreement with manual measurements, indicated by intraclass correlation coefficients (ICCs) of 0.975.

This study demonstrates advanced neural networks' ability to offer cost-effective, accurate, and scalable IOOA assessments using simple photographs [35]. Liu et al. [36] conducted a prospective cross-sectional study on 132 patients with intermittent exotropia, using a support vector machine (SVM) to determine the optimal timing for surgical interventions. The study utilized preoperative factors like deviation angle, binocular vision, and stereoacuity to predict the best moment for surgery. Post a 6-month follow-up, the SVM achieved an 82.1% accuracy rate, with an overall success rate of 63.6%, identifying the initial postoperative deviation angle as a crucial predictor of successful alignment. This research highlights the effectiveness of SVM in enhancing surgical outcomes for exotropia through precise timing based on detailed clinical data [36].

The aforementioned study illustrate that AI integration refines surgical planning and prognostication. By leveraging these technologies, ophthalmologists can achieve significantly improved accuracy in surgical outcomes, enhancing patient-specific treatment plans. Notably, these models facilitate the analysis of deviations, muscle actions, and postoperative alignments, allowing for highly individualized surgical interventions. This marks a significant shift in enhancing surgical precision and redefining strabismus management standards.

Discussion

The integration of artificial intelligence (AI) in strabismus management represents a significant advancement in ophthalmology, enhancing diagnostic precision and therapeutic efficacy. As AI technologies have begun to permeate every aspect of healthcare, from data collection and integration [41–45] to analysis and treatment [46], they demonstrate a potent synergy with medical professionals, showcasing their immense potential to enhance human health [47]. The advent of large models like ChatGPT has marked a pivotal shift, increasingly evident in the fourth industrial revolution, driving transformative changes in how healthcare is delivered.

While AI's application in healthcare, particularly in ophthalmology, is promising, it faces challenges such as the need for improved data-sharing mechanisms, validation, evaluation, and regulatory frameworks. Despite the impressive achievements of AI in enhancing strabismus screening, diagnosis, and treatment, there are significant limitations regarding the representativeness and generalizability of these advancements in real-world clinical settings. Current research often excludes numerous strabismus types, such as vertical, restrictive, sensory, paralytic strabismus, myasthenia gravis, nystagmus, and Duane syndrome, focusing primarily on pediatric populations and concomitant strabismus patients. This oversight highlights a critical gap, as there is a pressing

need for comprehensive screening and precise diagnostic processes that encompass all age groups and subtypes of strabismus. Thus, the utility of these AI models in diverse clinical environments remains questionable. Meanwhile, despite early advantages in AI research facilitated by the digitization in ophthalmology, which amassed extensive imaging and pathological data, the comprehensive integration of multimodal patient data remains scarce, and the reliance on single data sources for AI diagnostics and prognostic assessments presents substantial challenges. Typical methodologies primarily use ocular position photos or eye movement videos for diagnosing strabismus, or electronic health records for predicting surgical outcomes, without considering additional data inputs, such as the sensory exam results, the refractive status, the CT or MRI images, forced duction test, etc. Conversely, clinicians often employ a multimodal data approach, integrating those concurrent data streams to enhance the accuracy of disease diagnosis, prognostic evaluations, and therapeutic strategies. Current AI systems process static, singular time-point data, which fails to capture the dynamic, evolving nature of diseases and overlooks the rich historical context of patient data that could significantly refine predictive accuracies. Theoretically, as AI technologies advance, models should be capable of harnessing all available data sources, including those traditionally inaccessible to clinicians, such as genomics [48]. Moreover, in regions such as China, many hospitals still rely on non-digital, unstructured data collection methods due to a lack of digital resources and personnel, leading to incomplete or inaccurate medical data and hindering widespread AI research. Additionally, the field suffers from a scarcity of AI-driven strabismus management products that extend beyond diagnosis to include essential services such as patient referral and long-term follow-up, which are vital for ensuring sustained care and systematic data collection across various modalities and time points.

Moving forward, research utilizing deep learning in strabismus should embrace a more inclusive approach, incorporating patients of all ages, ethnicities, and all types of the condition. However, the diversity in irrelevant features such as eyelid size, eye color, and skin tone across different age groups and ethnic backgrounds poses additional challenges for model training. Complications such as eyelid ptosis, which require physical assistance during photo acquisition, can disrupt the accurate extraction of features by AI models. Additionally, ensuring privacy protection is critical as AI technologies handle sensitive medical data. Addressing these challenges is essential for maximizing the clinical utility of AI technologies. Future initiatives should focus on easily collectible multimodal data for strabismus research, integrating biosensors, electronic medical records, eye movement videos, ocular photos, and socio-environmental factors throughout the patient care timeline. Large AI models, such as Transformer-based architectures, have the potential to handle the complexity and scale of

multimodal data, offering enhanced generalization and predictive power. This integration not only facilitates extensive research but also enables the construction of ophthalmological knowledge graphs and the provision of personalized treatment recommendations, which are becoming increasingly feasible. Meanwhile, the development of an integrated AI-based strabismus management platform that facilitates paperless, comprehensive management of strabismus patients from screening through treatment, and enables the integration and collection of cross-modal patient information, will lay a solid foundation for a specialized, large-scale AI model in strabismus care. The platform should support AI applications in the screening, diagnosis, surgical parameter estimation, and prognosis prediction of strabismus, as well as intelligent follow-up. By leveraging a full-process, multimodal, and multi-timepoint intelligent platform, ophthalmologists can achieve more accurate, personalized, and effective strabismus management, ultimately improving patient outcomes and advancing the field of ophthalmology.

Conclusion

In conclusion, this review has highlighted significant advancements and emerging challenges in integrating AI into strabismus management. AI is revolutionizing screening, diagnosis, and surgical planning through cutting-edge technologies such as deep learning, enhancing accuracy and optimizing outcomes. However, critical gaps are evident, particularly the underrepresentation of diverse strabismus types and age groups in AI studies. These limitations underscore the necessity for AI systems that accommodate broader demographics and clinical scenarios. Additionally, the reliance on single-source

data in AI models, compared to the multimodal approach used in clinical settings, reveals an urgent need for integrating more comprehensive data sources, including longitudinal patient data. Addressing these challenges will be pivotal in advancing AI technology in ophthalmology, enhancing patient outcomes, and setting new standards in healthcare delivery.

Author contributions

DW and XH: Writing—original draft, Writing—review and editing, Methodology, Conceptualization, Data curation, Investigation. LC and PH: Writing—original draft, Data curation. GY: Investigation, Methodology, Resources, Supervision, Writing—original draft, Writing—review and editing, Conceptualization. LL: Supervision, Conceptualization, Critical revision of the manuscript. All authors contributed to the article and approved the submitted version.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by Project for disciplines of excellence for West China Hospital, Sichuan University, China (No. ZYJC23001).

Conflict of interest

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

1. Repka MX, Lum F, Burugapalli B. Strabismus, strabismus surgery, and reoperation rate in the United States: analysis from the IRIS registry. *Ophthalmology* (2018) **125**:1646–53. doi:10.1016/j.ophtha.2018.04.024
2. Mojon-Azzi SM, Mojon DS. Strabismus and employment: the opinion of headhunters. *Acta Ophthalmologica* (2009) **87**:784–8. doi:10.1111/j.1755-3768.2008.01352.x
3. Durnian JM, Noonan CP, Marsh IB. The psychosocial effects of adult strabismus: a review. *Br J Ophthalmol* (2011) **95**:450–3. doi:10.1136/bjo.2010.188425
4. Uretmen O, Egrilmez S, Kose S, Pamukcu K, Akkin C, Palamar M. Negative social bias against children with strabismus. *Acta Ophthalmologica Scand* (2003) **81**: 138–42. doi:10.1034/j.1600-0420.2003.00024.x
5. Cotter SA, Tarczy-Hornoch K, Song E, Lin J, Borchert M, Azen SP, et al. Fixation preference and visual acuity testing in a population-based cohort of preschool children with amblyopia risk factors. *Ophthalmology* (2009) **116**: 145–53. doi:10.1016/j.ophtha.2008.08.031
6. Wang X, Gao X, Xiao M, Tang L, Wei X, Zeng J, et al. Effectiveness of strabismus surgery on the health-related quality of life assessment of children with intermittent exotropia and their parents: a randomized clinical trial. *J Am Assoc Pediatr Ophthalmol Strabismus* (2015) **19**:298–303. doi:10.1016/j.jaapos.2015.04.007
7. Ambrósio R, Jr, Salomão MQ, Barros L, da Fonseca Filho JBR, Guedes J, Neto A, et al. Multimodal diagnostics for keratoconus and ectatic corneal diseases: a paradigm shift. *Eye Vis* (2023) **10**:45. doi:10.1186/s40662-023-00363-0
8. Zhou Y, Li G, Li H. Automatic cataract classification using deep neural network with discrete state transition. *IEEE Trans Med Imaging* (2020) **39**:436–46. doi:10.1109/tmi.2019.2928229
9. Li B, Chen H, Yu W, Zhang M, Lu F, Ma J, et al. The performance of a deep learning system in assisting junior ophthalmologists in diagnosing 13 major fundus diseases: a prospective multi-center clinical trial. *NPJ Digit Med* (2024) **7**:8. doi:10.1038/s41746-023-00991-9
10. Dong L, He W, Zhang R, Ge Z, Wang YX, Zhou J, et al. Artificial intelligence for screening of multiple retinal and optic nerve diseases. *JAMA Netw Open* (2022) **5**:e229960. doi:10.1001/jamanetworkopen.2022.9960
11. Dave T, Athaluri SA, Singh S. ChatGPT in medicine: an overview of its applications, advantages, limitations, future prospects, and ethical considerations. *Front Artif Intell* (2023) **6**:1169595. doi:10.3389/frai.2023.1169595
12. Harris E. Large Language models answer medical questions accurately, but can't match clinicians' knowledge. *Jama* (2023) **330**:792–4. doi:10.1001/jama.2023.14311
13. Singhal K, Azizi S, Tu T, Mahdavi SS, Wei J, Chung HW, et al. Large language models encode clinical knowledge. *Nature* (2023) **620**:172–80. doi:10.1038/s41586-023-06291-2

14. Long EP, Liu ZZ, Xiang YF, Xu AD, Huang JL, Huang XC, et al. Discrimination of the behavioural dynamics of visually impaired infants via deep learning. *Nat Biomed Eng* (2019) **3**:860–9. doi:10.1038/s41551-019-0461-9
15. Chen W, Li R, Yu Q, Xu A, Feng Y, Wang R, et al. Early detection of visual impairment in young children using a smartphone-based deep learning system. *Nat Med* (2023) **29**:493–503. doi:10.1038/s41591-022-02180-9
16. Miao Y, Jeon JY, Park G, Park SW, Heo H. Virtual reality-based measurement of ocular deviation in strabismus. *Computer Methods Programs Biomed* (2020) **185**: 105132. doi:10.1016/j.cmpb.2019.105132
17. Saisara U, Boonbrahm P, Chaiwiriya A. Strabismus screening by eye tracker and games. In: 14th international joint conference on computer science and software engineering (JCSSE). Nakhon Si Thammarat, Thailand, July 12–14, 2017 (IEEE) (2017).
18. Chen ZH, Fu H, Lo WL, Chi ZR. Eye-Tracking aided digital system for strabismus diagnosis. In: IEEE international conference on systems, man, and cybernetics (SMC), Hong Kong, Peoples R China, October 9–12, 2015 (Hong Kong, China: City Univ Hong Kong, IEEE Computer Soc) (2015). p. 2305–9.
19. Valente TLA, de Almeida JDS, Silva AC, Teixeira JAM, Gattass M. Automatic diagnosis of strabismus in digital videos through cover test. *Computer Methods Programs Biomed* (2017) **140**:295–305. doi:10.1016/j.cmpb.2017.01.002
20. Chen ZH, Fu H, Lo WL, Chi ZR. Strabismus recognition using eye-tracking data and convolutional neural networks. *J Healthc Eng* (2018) **2018**:1–9. doi:10.1155/2018/7692198
21. Ma SX, Guan YQ, Yuan YZ, Tai Y, Wang T. A one-step, streamlined children's vision screening solution based on smartphone imaging for resource-limited areas: design and preliminary field evaluation. *JMIR mHealth uHealth* (2020) **8**:e18226. doi:10.2196/18226
22. Kang YC, Yang HK, Kim YJ, Hwang JM, Kim KG. Automated mathematical algorithm for quantitative measurement of strabismus based on photographs of nine cardinal gaze positions. *Biomed Res Int* (2022) **2022**:1–7. doi:10.1155/2022/9840494
23. Sousa de Almeida JD, Silva AC, Teixeira JAM, Paiva AC, Gattass M. Computer-aided methodology for syndromic strabismus diagnosis. *J Digit Imaging* (2015) **28**:462–73. doi:10.1007/s10278-014-9758-0
24. de Oliveira ST, Carvalho SJ, Sousa de Almeida JD, Correa Silva A, Cardoso de Paiva A. Automatic ocular alignment evaluation for strabismus detection using U-net and ResNet networks. In: 2019 8th Brazilian conference on intelligent systems (BRACIS) proceedings, Salvador, Brazil, October 15–18, 2019, (2019). p. 239–44.
25. Mesquita M, Valente TLA, de Almeida JDS, Teixeira JAM, Medina FM, dos Santos AM. A mhealth application for automated detection and diagnosis of strabismus. *Int J Med Inform* (2021) **153**:6. doi:10.1016/j.ijmedinf.2021.104527
26. De Figueiredo LA, Dias JVP, Polati M, Carricondo PC, Debert I. Strabismus and artificial intelligence app: optimizing diagnostic and accuracy. *Translational Vis Sci and Technol* (2021) **10**:22. doi:10.1167/tvst.10.7.22
27. Huang X, Lee SJ, Kim CZ, Choi SH. An automatic screening method for strabismus detection based on image processing. *PLoS One* (2021) **16**:e0255643. doi:10.1371/journal.pone.0255643
28. Zheng C, Yao Q, Lu JW, Xie XL, Lin SB, Wang ZL, et al. Detection of referable horizontal strabismus in Children's primary gaze photographs using deep learning. *Translational Vis Sci and Technol* (2021) **10**:33. doi:10.1167/tvst.10.1.33
29. Mao KL, Yang YH, Guo C, Zhu Y, Chen C, Chen JC, et al. An artificial intelligence platform for the diagnosis and surgical planning of strabismus using corneal light-reflection photos. *Ann Transl Med* (2021) **9**:374. doi:10.21037/atm-20-5442
30. Wu D, Li Y, Zhang H, Yang X, Mao Y, Chen B, et al. An artificial intelligence platform for the screening and managing of strabismus. *Eye (Lond)* (2024) **38**: 3101–7. doi:10.1038/s41433-024-03228-5
31. Kim D, Joo J, Zhu G, Seo J, Ha J, Kim SC. Ieee. Strabismus classification using convolutional neural networks. In: 3rd international conference on artificial intelligence in information and communication (IEEE ICAIIC), Jeju Island, South Korea, April 20–23, 2021 (IEEE) (2021). p. 216–8.
32. Almeida JDSd, Silva AC, Teixeira JAM, Paiva AC, Gattass M. Surgical planning for horizontal strabismus using Support Vector Regression. *Comput Biol Med* (2015) **63**:178–86. doi:10.1016/j.cmpbiomed.2015.05.025
33. Leite FHF, Almeida JDSd, Cruz LBd, Teixeira JAM, Junior GB, Silva AC, et al. Surgical planning of horizontal strabismus using multiple output regression tree. *Comput Biol Med* (2021) **134**:104493. doi:10.1016/j.compbimed.2021.104493
34. Tang R, Wang W, Meng Q, Liang S, Miao Z, Guo L, et al. A strabismus surgery parameter design model with WGAN-GP data enhancement method. *J Phys Conf Ser (UK)* (2022) **2179**:012009. doi:10.1088/1742-6596/2179/1/012009
35. Lou LX, Huang XR, Sun YM, Cao J, Wang YQ, Zhang QN, et al. Automated photographic analysis of inferior oblique overaction based on deep learning. *Quant Imaging Med Surg* (2023) **13**:329–38. doi:10.21037/qims-22-467
36. Liu YL, Liu CG, Zhang W, Chen X, Zhao KX. Model of a support vector machine to assess the functional cure for surgery of intermittent exotropia. *Sci Rep* (2019) **9**(7):8321. doi:10.1038/s41598-019-38969-x
37. Smola AJ, Schölkopf B. A tutorial on support vector regression. *Stat Comput* (2004) **14**:199–222. doi:10.1023/b:stco.0000035301.49549.88
38. Borchani H, Varando G, Bielza C, Larrañaga P. A survey on multi-output regression. *WIREs Data Mining Knowledge Discov* (2015) **5**:216–33. doi:10.1002/widm.1157
39. Gu JX, Wang ZH, Kuen J, Ma LY, Shahroudy A, Shuai B, et al. Recent advances in convolutional neural networks. *Pattern Recognition* (2018) **77**:354–77. doi:10.1016/j.patcog.2017.10.013
40. Boxuan Y, Junwei F, Jun L. Residual recurrent neural networks for learning sequential representations. *Information (Switzerland)* (2018) **9**(56):14. doi:10.3390/info9030056
41. Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, et al. The UK Biobank resource with deep phenotyping and genomic data. *Nature* (2018) **562**: 203–9. doi:10.1038/s41586-018-0579-z
42. Chen ZM, Chen JS, Collins R, Guo Y, Peto R, Wu F, et al. China Kadoorie Biobank of 0.5 million people: survey methods, baseline characteristics and long-term follow-up. *Int J Epidemiol* (2011) **40**:1652–66. doi:10.1093/ije/dyr120
43. Nagai A, Hirata M, Kamatani Y, Muto K, Matsuda K, Kiyohara Y, et al. Overview of the BioBank Japan Project: study design and profile. *J Epidemiol* (2017) **27**:S2–S8. doi:10.1016/j.je.2016.12.005
44. Gaziano JM, Concato J, Brophy M, Fiore L, Pyarajan S, Breeling J, et al. Million Veteran Program: a mega-biobank to study genetic influences on health and disease. *J Clin Epidemiol* (2016) **70**:214–23. doi:10.1016/j.jclinepi.2015.09.016
45. Saenz A, Chen E, Marklund H, Rajpurkar P. The MAIDA initiative: establishing a framework for global medical-imaging data sharing. *The Lancet Digital Health* (2024) **6**:e6–e8. doi:10.1016/s2589-7500(23)00222-4
46. Bo ZH, Guo YC, Lyu J, Liang HR, He JX, Deng SJ, et al. Relay learning: a physically secure framework for clinical multi-site deep learning. *Npj Digit Med* (2023) **6**:204. doi:10.1038/s41746-023-00934-4
47. Dvijotham K, Winkens J, Barsbey M, Ghaisas S, Stanforth R, Pawlowski N, et al. Enhancing the reliability and accuracy of AI-enabled diagnosis via complementarity-driven deferral to clinicians. *Nat Med* (2023) **14**. doi:10.1038/s41591-023-02437-x
48. Acosta JN, Falcone GJ, Rajpurkar P, Topol EJ. Multimodal biomedical AI. *Nat Med* (2022) **28**:1773–84. doi:10.1038/s41591-022-01981-2