nature portfolio

Corresponding author(s): Tien Yin Wong, Bin Sheng, Xiaodong Sun

Last updated by author(s): Jun 21, 2026

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒ A description of all covariates tested
☐	☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	No special software was used for data collection.
Data analysis	Python version 3.8.20 (Python Software Foundation, Delaware, United States) was used for all computational experiments in this study. The following third-party Python packages were used: PyTorch version 1.10.0 and torchvision version 0.11.1 (Facebook, Menlo Park, California, United States) for deep learning model development; Pandas version 1.3.4 (The Pandas Development Team, United States) for data processing; and Scikit-learn version 1.3.2 (David Cournapeau, California, United States) for calculating accuracy. The code being used in the current study for developing the algorithm is provided at https://github.com/zycl2001/Retina4IRD.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

Data can only be shared for non-commercial academic purposes and will require a formal material transfer agreement. Interested investigators can obtain and certify the data transfer agreement and submit requests to X.S. (xdsun@sjtu.edu.cn) Investigators who consent to the terms of the data transfer agreement, including, but not limited to, the use of these data only for academic purposes, and to protect the confidentiality of the data and limit the possibility of identification of patients, will be granted access. Requests will be evaluated on a case-by-case basis within one month before receipt of a response. All data shared will be deidentified. For the reproduction of our algorithm code, we have also deposited a minimum dataset at Zenodo (https://zenodo.org/records/20489924), which is publicly available for scientific research and noncommercial use. Source data are provided with this paper.

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	This study strictly adheres to the SAGER guidelines. Consistent with these recommendations, we have applied the term "sex" solely in reference to biological attributes. Participant sex was determined from official identity (ID) documentation, and the data were classified into two categories: Female and Male. Descriptive data analysis was reported by sex is described in Table 1 . Post hoc subgroup analyses according to sex are presented in Figure 5.
Reporting on race, ethnicity, or other socially relevant groupings	For the development and validation of Retina4IRD, multi-ethnic datasets were used. The cohort demographics from different ethnics were shown in Table 1. The performance of Retina4IRD in multi-ethnic datasets were presented in Figure 3, Figure 5, Extended Data Figure 1, and Extended Data Figure 2.
Population characteristics	Detailed population characteristics was given in Table 1.
Recruitment	This is a two-stage study. In Stage 1, we retrospectively collected 5,167 images and 8,901 metadata items from 1,843 genetically confirmed individuals across China, South Korea, and Poland. Multi-ethnic but predominantly East Asian sample may introduce geographic bias. Stage 2 is a randomized controlled validation trial including 300 participants conducted in China to evaluate the effectiveness of Retina4IRD in real-world clinical settings. Self-selection bias may overestimate effectiveness, as volunteered physicians and patients may have stronger interest in AI.
Ethics oversight	This study was conducted in accordance with all applicable ethical regulations. Stage 1, a retrospective model development study, was approved by the Ethics Committee of the National Clinical Research Centre for Eye Diseases (2022KY140, approved December 30, 2022; 2022KY140-2, approved December 29, 2023). Stage 2, a randomized controlled trial, was approved by the same Ethics Committee (2024-451, approved November 6, 2024) and registered with ClinicalTrials.gov (NCT06839170).

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	In the RCT, a sample size of 300 for statistical power above 95% with a two side Alpha = 0.05, determining based on the hypothesis that the AI-based CDSS group would achieve 85% accuracy in top-5 gene prediction, compared to 60% in the specialist-only group.
Data exclusions	During screening, 107 cases were excluded (88 declined, 19 met exclusion criteria). In the final analysis, 5 additional participants were excluded for missing NGS test reports. All exclusion criteria were pre-established and applied consistently. Exclusion criteria for the RCT were refusal to undergo NGS genetic testing, fundus photography, or OCT examination; presence of other ocular conditions that may interfere with retinal examination; history of intraocular surgery in either eye within 6 months prior to screening; systemic severe diseases, intellectual disability, or psychiatric disorders; participation in other interventional clinical trials, such as stem cell or gene therapy trials.
Replication	Replication is not relevant. We used independent validation cohorts and experiments to test Retina4IRD, and it achieved similar performance.

Randomization	Randomization was performed in a 1:1 ratio using a computer-generated pseudo-random number sequence.
Blinding	Participated patients and assessors will be masked.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☐	☒ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Clinical data

Policy information about clinical studies

All manuscripts should comply with the ICMJE guidelines for publication of clinical research and a completed CONSORT checklist must be included with all submissions.

Clinical trial registration	ClinicalTrials.gov NCT06839170
Study protocol	Protocol was attached in Supplementary.
Data collection	The trial was conducted from February 27, 2025 to July 1, 2025 at Shanghai General Hospital; Eye & ENT Hospital of Fudan University; Beijing Tongren Hospital; Zhongshan Ophthalmic Center, Sun Yat-sen University; West China Hospital, Sichuan University; Henan Provincial People's Hospital; People's Hospital of Ningxia Hui Autonomous Region. Collected data include color fundus photo, OCT image, sex, age, family history, age onset, disease duration and genotype.
Outcomes	The primary outcome measure was top-5 gene prediction accuracy, defined as the proportion of cases in which the correct genotype was ranked among the top 5 predictions. Specifically, a ranked list of candidate genotypes was generated for each case by the two arms. A prediction was considered correct if the actual genotype, as confirmed by the subsequent NGS report (the gold standard), fell within the top-5 ranked predictions. The secondary outcomes included the accuracy of top-1 to top-4 gene predictions, defined as the proportion of cases in which the correct genotype prediction was ranked among the top-N predictions. A prediction was considered correct if the actual genotype, as confirmed by the subsequent NGS report, fell within the top-N ranked predictions.

Plants

Seed stocks	not applicable
Novel plant genotypes	not applicable
Authentication	not applicable