

# The German National Cohort: Ophthalmic Assessment, Baseline Profile, and Potential for Artificial Intelligence–Based Eye Research

Camila Roa, MSc,<sup>1,2,\*</sup> Ansgar Beuse, MD,<sup>3,\*</sup> Alexandra Schweig, MD,<sup>4</sup> Sarah Müller, MSc,<sup>1,2</sup> Klaus Berger, MD, MSc,<sup>5</sup> Caroline Brandl, MD,<sup>6,7</sup> Titus Brinker, MD,<sup>8</sup> Anne Elbrecht, MD,<sup>9</sup> Robert Finger, MD, PhD,<sup>10</sup> Gerd Geerling, MD,<sup>11</sup> Karin Halina Greiser, MD, PhD,<sup>8</sup> Carsten Grohmann, MD, PhD,<sup>3</sup> Kathrin Günther, PhD,<sup>9</sup> Iris Heid, PhD,<sup>6</sup> André Karch, MD, MSc,<sup>5</sup> Thomas Keil, MD, MSc,<sup>12,13,14</sup> Jessica Krepel, PhD,<sup>15</sup> Michael Leitzmann, MD, MPH,<sup>16</sup> Claudia Meinke-Franze, PhD,<sup>17</sup> Annette Peters, PhD, MPH,<sup>18,19</sup> Sabine Schipf, PhD,<sup>17</sup> Matthias Schulze, DrPH,<sup>20,21</sup> Alexander K. Schuster, MD, MSc,<sup>22</sup> Stefan N. Willich, MD, MPH,<sup>11</sup> Martin A. Leitritz, MD,<sup>4</sup> Marius Ueffing, PhD,<sup>23,\*\*</sup> Philipp Berens, PhD<sup>1,2,\*\*</sup>

**Objective:** To describe the ophthalmic examination protocol within the German National Cohort (NAKO)/NAKO Gesundheitsstudie, to report the baseline profile of participants undergoing ophthalmic assessment, and to illustrate the potential as a population-based resource for artificial intelligence (AI) research in eye health.

**Design:** Baseline analysis of ophthalmic data within the nationwide, population-based multicenter prospective NAKO study.

**Participants:** Forty-eight thousand four hundred and sixty adults in the ophthalmic level 2 module of 205 053 adults enrolled in NAKO, aged 19–74 years, with a mean age  $48.9 \pm 12.5$  years, and 52.7% were male.

**Methods:** All participants underwent standardized biomedical examinations and detailed questionnaire-based data collection, including nondilated color fundus imaging, visual acuity testing, and a brief ocular history. Ocular health was summarized using descriptive statistics. Fundus image quality and morphological features (e.g., cup-to-disc ratio and arteriolar-to-venular ratio) were assessed using open-source deep learning models. Standard deep learning architectures were trained on the fundus images to predict age, sex, and blood pressure.

**Main Outcome Measures:** Percentage of fundus images graded as good quality; mean absolute error (MAE) for age and blood pressure prediction; accuracy for sex prediction.

**Results:** The analysis includes 48 460 participants who completed the ophthalmic baseline examination of NAKO. Mean visual acuity (logarithm of the minimum angle of resolution) was  $0.01 \pm 0.20$  (left eye, mean  $\pm$  standard deviation) and  $0.03 \pm 0.21$  (right eye). Self-reported ocular disease prevalence was 4.2% for cataract, 2.0% for glaucoma, and 0.9% for macular degeneration; 68.2% of fundus images were classified as gradable using 4 deep learning–based quality grading models. Morphological features showed systematic differences across age groups. Standard deep learning architectures showed comparative performance to the state-of-the-art for age, sex, and blood pressure prediction (2.96 MAE for age prediction, 0.84 accuracy for sex prediction, and 10.78 and 6.97 MAE for systolic and diastolic blood pressure prediction).

**Conclusions:** The German National Cohort provides a large-scale, nationwide population-based resource with visual acuity measurements and systemic health indicators, as well as color fundus images in about 50 000 NAKO participants. These data lay the groundwork for studying eye health in the general adult population in Germany and can serve as a strong foundation for developing and validating AI tools in eye health research.

**Financial Disclosure(s):** Proprietary or commercial disclosure may be found in the Footnotes and Disclosures at the end of this article. *Ophthalmology Science* 2026;6:101408 © 2026 American Academy of Ophthalmology, Inc. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).



Supplemental material available at [www.ophtalmologyscience.org](http://www.ophtalmologyscience.org).

The German National Cohort (NAKO)/NAKO Gesundheitsstudie is a nationwide, population-based prospective study designed to investigate the determinants, risk factors, early markers, trajectories, and long-term consequences of major chronic diseases in Germany.<sup>1,2</sup> With 205 053 participants enrolled at baseline across 18 study centers, NAKO is one of the largest deeply phenotyped cohorts worldwide. It is among large population-based studies with >100 000 participants, including nationwide studies such as the UK Biobank (UKB)<sup>3</sup> and the China Kadoorie Biobank.<sup>4</sup> The German National Cohort's comprehensive design incorporates multimodal imaging, biospecimen collection, environmental and behavioral profiling, and a planned overall duration of 25 to 30 years.<sup>1,5</sup> A subgroup of >55 000 participants was offered an extended examination program (so called level 2 examinations).

Within this framework, the ophthalmology module was established as one of the level 2 investigations, designed to serve 3 purposes that define its intended scope of use. First, NAKO provides population-based reference data on visual acuity and self-reported ocular disease prevalence in the German adult population, enabling epidemiological studies of eye health across age, sex, and region, and—as follow-up data accumulate—longitudinal assessment of visual decline and its determinants. Second, the standardized acquisition of color fundus images across 18 study centers creates a large-scale resource for the development, training, and external validation of deep learning models for automated retinal phenotyping and ophthalmic research, linking retinal features to the extensive systemic health data collected in the NAKO. Third, the combination of imaging, visual function, and self-reported ocular disease in a deeply phenotyped cohort with a planned follow-up of 25–30 years provides a foundation for investigating associations between ocular measures and nonocular conditions, such as diabetes, hypertension, and cardiovascular disease—in both directions.

In this paper, we describe the examination protocol and quality assurance framework of the NAKO ophthalmology module, report the baseline profile of the 48 460 participants who underwent ophthalmic assessment, and demonstrate the suitability of the data for artificial intelligence (AI)-based research by evaluating fundus image quality with open-source deep learning models, extracting morphological features, and training standard architectures to predict age, sex, and blood pressure. This paper serves as the primary methodological reference for the ophthalmic assessment within the NAKO. Researchers using NAKO ophthalmic data—including fundus images, visual acuity measurements, derived quality gradings, morphological features, or trained models—are asked to cite this paper alongside the general NAKO study reference.<sup>1</sup>

## Methods

### Design, Baseline Recruitment, and Ethical Approval

The NAKO is a large, prospective, population-based cohort study conducted in Germany. Participants were randomly selected from compulsory population registries in 16 German regions using age-

and sex-stratified sampling. They were examined between 2014 and 2019 in 18 study centers representing urban, suburban, and rural settings.<sup>6</sup> The NAKO aimed to recruit approximately 200 000 residents between the ages of 20 and 69.<sup>5</sup>

The baseline recruitment yielded 205 415 participants aged 19–74 years. The inclusion criteria were: residing in the study catchment area, being within the target age range at the time of invitation, ability to provide written informed consent, and proficiency in the German language to understand questionnaires and instructions. The exclusion criteria included being unable to consent. The overall response rate averaged 15.6%, with higher participation among older individuals and women with differences across regions.<sup>1,6</sup> Recruitment was monitored to ensure representativeness across predefined strata. After 362 participants withdrew informed consent, the sample consisted of 205 053 individuals as of October 2019.<sup>1</sup>

Baseline assessment included standardized, computer-assisted face-to-face interviews, self-administered questionnaires, standardized physical and medical examinations, and biosample collection. A core set of questionnaires and examinations was offered to all participants (so called level 1 modules), and an extended set of in-depth examinations and questionnaire modules (level 2 modules) was offered to a 20% random sample of all participants plus to additional participants undergoing magnetic resonance imaging at 5 of the 18 study centers. Information about the items in the questionnaire is available as part of the NAKO Data Handbook<sup>1</sup>. For training of the AI-based models regarding estimation of age, sex, and blood pressure, we used data obtained during interview and during measurement of the sitting blood pressure with the OMRON-HEM 705 IT.<sup>1</sup>

The study protocol of NAKO was approved by the ethics committees of all participating institutions, and all participants provided written informed consent in accordance with German legal and data protection requirements. It is conducted in accordance with the Declaration of Helsinki and national standards for good clinical and epidemiological practice.<sup>1</sup>

### The Ophthalmology Module

The intensified level 2 examination of the NAKO includes an ophthalmology module comprising visual acuity testing, non-mydriatic color fundus imaging, and a standardized self-report questionnaire on ocular disease history. The module was conducted in 56 971 participants of the baseline examination (Fig 1). Of those, 48 460 completed at least one visual acuity measurement and one fundus image for  $\geq 1$  eye, according to the data set provided by the TransferHub (Fig 1, Table 1). The complete eye examination required approximately 10 minutes.

**Visual Acuity Assessment.** Visual acuity was tested monocularly while participants wore their walk-in distance correction (glasses or contact lenses). An additional occluder fully covered the nontested eye and incorporated a central aperture in front of the eye under examination. If visual acuity exceeded a logarithm of the minimum angle of resolution value of 0.0, a rotatable insert with multiple 2.0 mm perforated openings was additionally positioned in front of this aperture to enhance letter recognition and mitigate residual refractive errors. Illuminated optotype charts with Latin letters, based on the ETDRS protocol, were used at a testing distance of 4 m. In case of room length restrictions, a fixed mirror was used to achieve the required distance. Visual acuity was assessable from  $-0.20$  to  $1.20$ . The test was started by reading all letters from the top of the letter chart from left to right. It was terminated when  $<3$  of 5 letters on the same line were correctly identified. Results—including the type of refractive correction, identified optotypes, and notable observations—were digitally documented by study staff.<sup>7</sup> The visual acuity protocol was

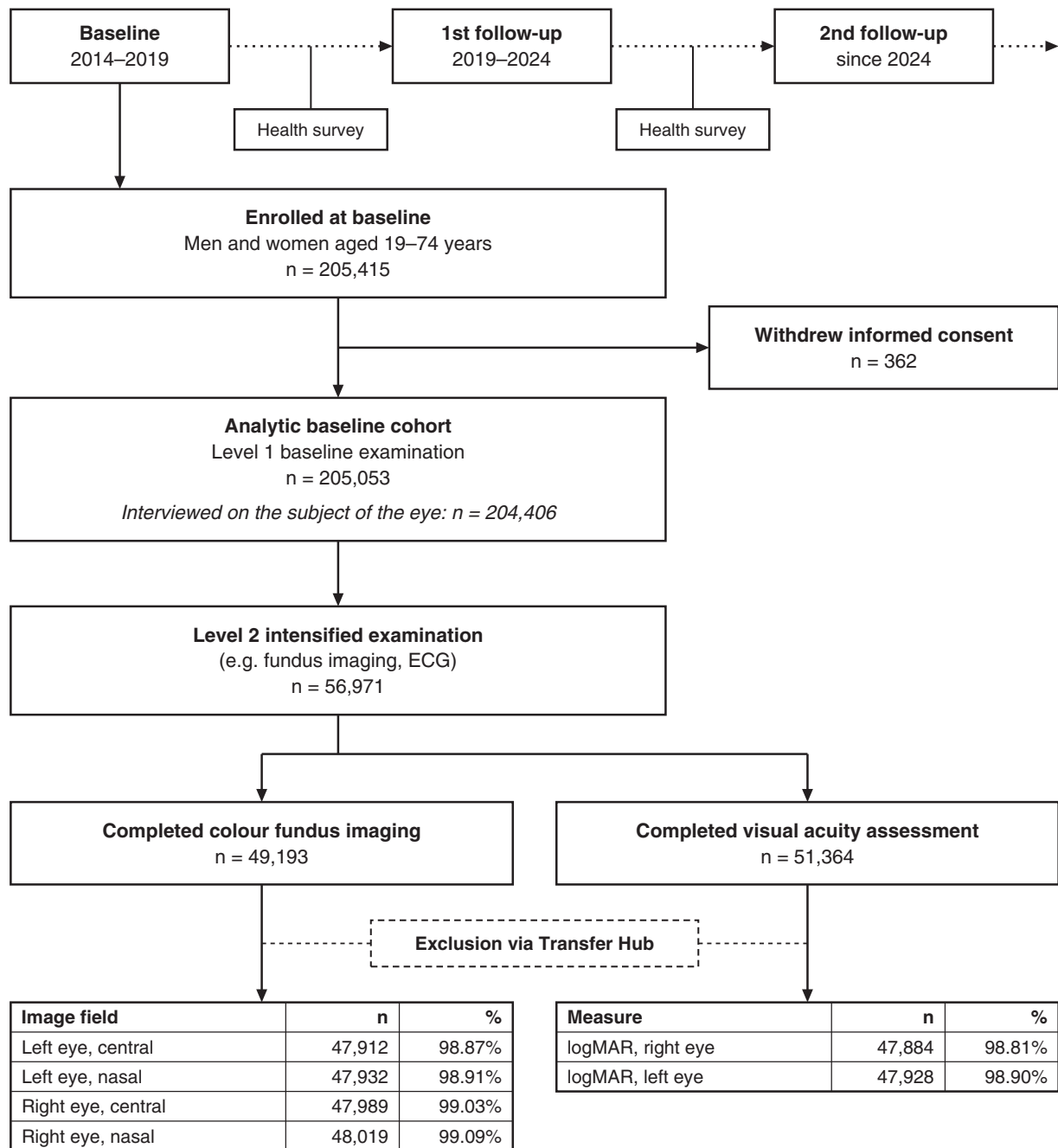


**Figure 1.** Flowchart of study participants for the baseline examination in the level 2 ophthalmology submodule of the NAKO. Numbers are reported as in n = participants. ECG = electrocardiography; logMAR = logarithm of the minimum angle of resolution; NAKO = German National Cohort.

harmonized with and adapted from the UKB examination protocol to permit future joint analyses.

**Fundus Imaging.** Retinal imaging was performed without mydriasis using a DRS fundus camera (Revenio group, icare, Finland; formerly CenterVue S.p.A.). The camera has a field of view of  $45^\circ \times 40^\circ$ , an autofocus range of  $-15$  to  $+15$  D, internal fixation light-emitting diodes, a working distance of 37 mm, and a 5-megapixel sensor ( $2592 \times 1944$  pixels), yielding a resolution of 48 pixels/degree and a retinal resolution of  $15 \mu\text{m}^2$  for the standard eye. For each eye, one macula-centered and one optic disc-centered image was acquired. If the first image was judged

to be of insufficient quality, the study staff repeated the acquisition; all images—including retakes—were saved. Fundus imaging was performed in a darkened room and caused only minor inconvenience (brief glare) for the participants.

**Ocular Disease History.** Alongside the eye examination, participants were assessed for self-reported macular degeneration, cataract, optic nerve inflammation, and glaucoma using a standardized questionnaire.

**Exclusion Criteria.** The exclusion criteria for visual acuity testing and fundus imaging were assessed prior to testing and included: participant refusal, severe ocular infection, ocular

Table 1. Summary of Demographic, Systemic, and Eye Health Variables of n = 48 460 Participants Provided by NAKO TransferHub from the Baseline in the NAKO Ophthalmology Data Set

| Variable                        | Total (n = 48 460)            |
|---------------------------------|-------------------------------|
| Continuous                      | mean $\pm$ SD [min, max]      |
| Age (years)                     | 48.91 $\pm$ 12.53 [19, 74]    |
| Systolic blood pressure (mmHg)  | 127.14 $\pm$ 16.16 [75, 236]  |
| Diastolic blood pressure (mmHg) | 78.39 $\pm$ 9.96 [44, 136]    |
| Left eye logMAR                 | 0.01 $\pm$ 0.20 [−0.20, 1.20] |
| Right eye logMAR                | 0.03 $\pm$ 0.21 [−0.20, 1.20] |
| Categorical                     | n (%)                         |
| Sex                             |                               |
| Male                            | 25 534 (52.69%)               |
| Female                          | 22 926 (47.31%)               |
| Smoking status                  |                               |
| Not a smoker                    | 22 836 (47.12%)               |
| Former smoker                   | 15 303 (31.58%)               |
| Smoker                          | 9185 (18.95%)                 |
| Missing                         | 1136 (2.34%)                  |
| High blood pressure (yes)       | 11 566 (23.87%)               |
| Diabetes mellitus (yes)         | 2380 (4.91%)                  |
| Cataracts (yes)                 | 2042 (4.21%)                  |
| Glaucoma (yes)                  | 971 (2.00%)                   |
| Macular degeneration (yes)      | 429 (0.89%)                   |

logMAR = logarithm of the minimum angle of resolution; NAKO = German National Cohort; SD = standard deviation.

Continuous variables are reported as mean  $\pm$  SD [min, max], and discrete variables as n (%). Note: Percentages are relative to the total number of participants. Missing values for disease variables were high blood pressure, 42 (0.09%); diabetes mellitus, 64 (0.13%); cataracts, 121 (0.25%); glaucoma, 151 (0.31%); macular degeneration, 161 (0.33%).

surgery within the prior 4 weeks, prosthetic eye, and known blindness. Fundus imaging was additionally contraindicated in participants with photosensitive epilepsy. Testing could be discontinued at any time by the participant, examiner, or due to technical issues (e.g., camera malfunction).

## Quality Assurance and Organizational Oversight

The ophthalmology module was developed by a panel of ophthalmic, epidemiological, and information-technology experts from Germany (NAKO eye expert group), which is responsible for the development of the ophthalmology program, oversight, and support of future project development. A pretest study was conducted to test standard operating procedures for visual acuity assessment and color fundus imaging as well as different fundus cameras conducted.<sup>7</sup> Standard operating procedures were modified to maximize information content versus participants' time in the study center to ensure feasibility at the scale of NAKO.

All examination procedures followed standard operating procedures, and all study personnel were trained and certified before collecting data. The NAKO-retina competence unit is responsible for overall management and quality assurance of the eye data across all 18 study centers. This includes regular (remote) site visits to ensure appropriate and equal on-site conditions, training sessions to ensure the integrity and accuracy of examinations and data processing, assessment of incidental findings, and continuous quality assessment through ongoing communication with the central data collection units and study centers.

Following the approval for pseudonymized analysis, fundus images were transferred to the Reading Centre at the Department

of Ophthalmology, University Hospital Tübingen, Germany. Manual grading and annotation of the images from a list of predefined findings was performed at the Reading Centre. These annotations are being made available through the NAKO TransferHub. The Reading Centre provides specialized personnel, state-of-the-art equipment, and expert knowledge to support computer-assisted, standardized evaluation of the data.

## Data Availability and Governance

All data described in this paper are accessible through the official NAKO TransferHub (<https://transfer.nako.de>). Access requires submission of a Use/Access Application, which is reviewed by the NAKO Use and Access Committee for scientific merit, feasibility, consistency with the study's objectives and participants' informed consent, and compliance with legal and ethical standards. Upon approval by the Board of Directors and conclusion of a Use/Access Agreement, applicants receive a time-limited, nonexclusive, nontransferable right to use the requested data in pseudonymized form for the approved purpose. Applicant institutions must be located within the European Economic Area or in a jurisdiction with adequate data protection safeguards. Commercial intentions must be disclosed and require separate approval. Users are required to submit results data and metadata to NAKO e.V. within 12 months of project completion and to delete all provided data upon expiry of the authorized period of use. Imaging data may only be used in publications in a manner that precludes reidentification of participants.

To protect participant privacy, the NAKO employs a multilevel pseudonymization architecture in accordance with the guidelines of the TMF (Technology, Methods, and Infrastructure for Networked Medical Research). Personally identifying data (e.g., name, address, and date of birth) are stored exclusively at the organizationally independent Trust Centre (Unabhängige Treuhandstelle) and are never accessible to researchers. Medical study data are stored and processed only under domain-specific pseudonyms—Independently generated random identifiers that differ across data systems (e.g., study database, laboratory information system, and research database). For each approved research project, a project-specific analysis data set is compiled with a further layer of repseudonymization, taking into account the reidentification risk arising from the specific combination of variables provided. Due to this governance framework and the requirements of German data protection law, the data cannot be deposited in an unrestricted public repository. The Use and Access Policy, Data Protection and IT Security Concept, and ethics documentation are available at <https://nako.de/en/research/transferhub/>. A Healthsheet for this data set, following the framework of Rostamzadeh et al, is provided as Supplementary Material.

## Statistical Analyses

We performed a descriptive analysis of the study population, summarizing basic demographic information (sex, age, and study site), general health related variables (blood pressure, smoking status, and diabetes), and ophthalmic variables (visual acuity, presence of eye diseases in self-report). Continuous variables are presented as means and standard deviations as well as minimum and maximum values. Categorical variables were summarized using absolute counts and percentages. Age-group distributions were calculated separately for men and women and reported as proportions of the total sample. Missing data were reported explicitly and no imputation was performed. All analyses were performed in Python (Python Software Foundation) using standard statistical procedures.

## Automated Fundus Image Analysis

**Preprocessing.** We used the Fundus Image Toolbox (FIT)<sup>8</sup> to crop and center the fundus images appropriately.

**Automatic Assessment of Image Quality.** To automatically assess the quality of the fundus images, we used 4 different open-source software packages that include quality prediction models. For all models, we used default parameters and applied them to the preprocessed fundus images resized to 512 × 512.

- Fundus Image Toolbox<sup>8</sup>: a python package for fundus image processing that includes quality grading, fovea and optic disc localization, vessel segmentation, registration, and circle cropping. Quality is rated on a scale from 0 to 1 (ungradable to gradable) by an ensemble of standard deep learning architectures (ResNets and EfficientNets) trained on the DeepDRiD and DrimDB data sets.<sup>9,10</sup> We used the default threshold of 0.25, classifying images with a score of  $\geq 0.25$  as gradable.
- AutoMorph<sup>11</sup>: a pipeline for automated analysis of retinal images that includes preprocessing, quality grading, vessel, artery, vein, and disc segmentation and morphological feature measurement. For quality grading, a standard deep learning architecture (EfficientNet-B4) trained on the EyeQ data set<sup>12</sup> classifies images as good, usable, or reject. We considered good and usable images as gradable.
- VascX<sup>13</sup>: a pipeline for fundus analysis that includes preprocessing, quality grading, vessel artery-vein and disc segmentation, and fovea localization. The quality grading model is an ensemble of fine-tuned deep learning models (ResNet101) on the EyeQ data set that classifies images as good, usable, or reject. We considered good and usable images as gradable.
- RetiRatey (<https://github.com/justinengelmann/RetiRatey>): a fine-tuned small convolutional variant of the MobileNet-V4 family that rates the quality of retinal images on a scale from 1 to 4 (1: Very good, 2: Good, 3: Borderline, and 4: Bad) for different regions of interest: the macula, the disc, and the vessels. We considered images with scores  $< 3$  on all 3 regions of interest as gradable.

We validated automatic grading procedures on a subset of  $n = 100$  images of different quality sampled randomly. These images were graded by 5 expert annotators as excellent, good, fair, poor, and unusable. The average Cohen  $\kappa$  between raters was high on average, indicating high interrater agreement ( $\kappa = 0.73$ ). The gradings of different annotators were aggregated by averaging and majority vote. Spearman  $\rho$  between gradings and continuous model outputs as well as Cohen between binarized gradings and model outputs was used to quantify association strength. Automatic gradings by all 4 methods will be made available via the NAKO TransferHub.

**Automatic Extraction of Morphological Fundus Image Features and Epidemiological Analysis.** We used the AutoMorph toolbox<sup>11</sup> to extract morphological measurements from the macula-centered retinal images, using its built-in quality prediction module as part of the end-to-end pipeline, which was designed and tuned for morphological feature extraction. We focused on the following measurements: the vertical cup to disk ratio, along with the central retinal artery equivalent (CRAE), the arteriolar-to-venular ratio (AVR), and the vessel density of zone b.<sup>14</sup> We did not correct for magnification due to differences in axial length, corneal curvature, or refractive error, as this information was not available in the NAKO module. For these variables, we analyzed their relationship with age and sex using generalized additive models via the R package mgcv (version 1.9-4), fitting a smooth term for age, which was allowed to differ between sexes,

resulting in the formula  $y \sim \text{sex} + s(\text{age}, \text{by} = \text{sex})$ , where  $y$  is the variable of interest.<sup>15</sup> We removed outliers using the interquartile range (IQR) method, applied independently to each feature. Outliers below  $Q1 - 1.5 \times \text{IQR}$  or above  $Q3 + 1.5 \times \text{IQR}$  were excluded, where  $\text{IQR} = Q3 - Q1$ . Omitting outlier removal did not qualitatively change the results. Extracted morphological properties will be made available through the NAKO TransferHub.

**Deep Learning–Based Prediction of Systemic Variables.** We trained several standard deep learning models—ResNet18, ResNet50, Inceptionv3, Xception, EfficientNetv2-Small, ViT-Small, and Swin-Tiny<sup>16–21</sup>—to predict age, blood pressure, and sex from fundus images from the NAKO baseline examination. The size of the transformer models was chosen to yield models of a size in the same order of magnitude as a ResNet50. Preliminary experiments revealed that ImageNet pretraining was helpful to obtain better performing models in less time. Thus, all models were initialized with ImageNet weights and fine-tuned for a maximum of 25 epochs with early stopping. We used a batch size of 256 and an AdamW optimizer with a learning rate of  $10^{-4}$  and weight decay of  $10^{-4}$ . We trained the classifiers with cross-entropy loss and the regression models with mean squared error loss. The blood pressure model was trained with 2 prediction heads as a multitask model for both systolic and diastolic blood pressure. All trained models will be made available through the NAKO TransferHub and are available from the authors with a valid NAKO application.

For the development of deep learning models, we used images that were graded as “good” by the FIT.<sup>8</sup> We always selected the highest-quality image available for each participant per eye and field. This resulted in a total of 153 134 images from 45 247 participants. The images were resized to 224 × 224 pixels and split into training set (90%) and test set (10%), ensuring that all images from a given participant appeared only in one split. All models were trained using fivefold cross-validation on the training split, with one of the folds serving as the validation set. We selected the best model based on validation performance metrics and then evaluated it on the test set using a bootstrap procedure with 1000 resamples of 15 000 images from the test set. We report the average test set performance and the 95% confidence intervals (CIs), computed as the 2.5 and 97.5 percentiles for various metrics, including accuracy and area under the receiver operating characteristic curve for sex and mean absolute error (MAE) and the coefficient of determination ( $R^2$ ) for age, systolic blood pressure, and diastolic blood pressure.

## Results

### Population Profile

The data set consisted of 48 460 participants who had taken part in the baseline ophthalmology module and had  $\geq 1$  fundus image for  $\geq 1$  eye and a visual acuity examination, as provided by the NAKO TransferHub (Fig 1). Of these, 25 534 were male (52.7%) and 22 926 were female (47.3%) (Table 1). Thus, the percentage of women receiving the baseline eye examination was lower than the percentage of women in the overall NAKO study (50.5%).<sup>22</sup> The mean age of the participants was  $48.9 \pm 12.5$  years, ranging from 19 to 74 years old, similar to the overall NAKO average ( $49.9 \pm 12.8$  years). The different age ranges were well represented: 10.3% of participants were between 20 and 29 years old; 11.3% were between 30 to 39 years old; 27.8% were between 40 and 49 years old; 27.4% were between 50 and 59 years old; 21.5% were between 60 and 69 years old; and 1.7%

of participants were between 70 and 79 years old (Supplementary Table 1, available at [www.ophtalmologyscience.org](http://www.ophtalmologyscience.org)). This made the cohort comparatively young at the start of the study.<sup>1,3</sup> Participants from all study centers completed the ophthalmic assessment (Supplementary Table 2, available at [www.ophtalmologyscience.org](http://www.ophtalmologyscience.org)).

The cardiovascular and metabolic profile of the participants was typical of a comparatively young adult population cohort (see Table 1; Supplementary Tables 3 and 4, available at [www.ophtalmologyscience.org](http://www.ophtalmologyscience.org)): The mean systolic blood pressure was  $127.1 \pm 16.2$  mmHg (mean  $\pm$  standard deviation, range 75–236) and the mean diastolic blood pressure was  $78.4 \pm 10.0$  mmHg (range 44–136), with slightly higher systolic blood pressure in men than in women. Based on self-reports, 23.9% of participants reported high blood pressure, while 76.0% reported no history of hypertension. Additionally, 19.4% reported to be active smokers, and 32.6% had smoked in the past, with fewer past smokers in women than in men. In total, 4.9% of participants reported to have diabetes mellitus, with less women than men.

The participants reported a variety of ocular diseases, including cataracts (4.2%), glaucoma (2.0%), and macular degeneration (0.9%) (Table 1). Visual acuity, assessed using logarithm of the minimum angle of resolution measurements, was on average  $0.01 \pm 0.20$  in the left eye (mean  $\pm$  standard deviation) and  $0.03 \pm 0.21$  in the right eye, with ranges from  $-0.20$  to  $1.20$  in both eyes, and similar values for men and women. Of the 48 460 participants, 203 493 fundus images were taken, 51 140 from the left central field, 51 112 from the left nasal field, 50 348 from the right central field, and 50 893 from the right nasal field. As per instruction in the standard operating procedures, images were retaken 1 to 2 times for some participants, if the first image was judged to be of insufficient quality; we included all these images in the data set.

## Quality of Fundus Images

We found that the acquired color fundus images were generally of high quality. However, there were also some images of insufficient quality for grading (Fig 2). To systematically evaluate the quality of the acquired fundus images and its dependency on age, we automatically analyzed all 203 493 fundus images of the 48 460 participants with 4 deep learning models for fundus image quality assessment (see Methods). We validated the automatic quality gradings on a subset of 100 images using 5 expert annotators. We found that the average expert quality grades correlated highly with the quality assessment by the deep learning models (Spearman  $\rho = 0.89/0.88/0.80/0.54$  for FIT, AutoMorph, VascX, and RetiRatey, respectively; Supplementary Table 5, available at [www.ophtalmologyscience.org](http://www.ophtalmologyscience.org)). Similar results were also obtained for majority voting for aggregating expert ratings as well as using Cohen  $\kappa$  between binarized model outputs and gradings.

Overall, the 4 algorithms, FIT, RetiRatey, AutoMorph, and VascX, graded 155 857 (76.6%), 177 172 (87.1%),

154 542 (75.9%), and 151 415 (74.4%) out of 203 493 images as gradable, and for 138 885 out of 203 493 (68.2%), all 4 algorithms agreed that these images were of high quality. This percentage is higher than that of other population-based cohort studies, such as the ophthalmic cohort of the UKB (high quality in agreement of all 4 algorithms: 48.7%)<sup>23</sup> assessed by us using the same algorithms and settings. Depending on the algorithm and the fundus image field, we found that between 70.1% and 91.4% of participants had  $\geq 1$  image for  $\geq 1$  eye that was classified as good quality (Table 2). Interestingly, the proportion of gradable images was typically higher for the central fields in both eyes, and image quality correlated with the order in which images were taken (right central, right nasal, left central, and left nasal). Image quality declined with increasing age (from 80% to 90% gradable images below 40 years to 40.8% at oldest age group; Supplementary Table 6, available at [www.ophtalmologyscience.org](http://www.ophtalmologyscience.org)) similarly in men and women.

## Epidemiological Analysis of Basic Morphological Fundus Properties

To demonstrate the potential of the data set for epidemiological analysis, we used an automatic, externally validated deep learning tool<sup>11</sup> to extract standardized morphological measurements related to the optic nerve and microvascular health of the fundus, such as the vertical cup-to-disc ratio, CRAE, AVR, and vessel density (see Methods). On average, vertical cup-to-disc ratio was 0.501, CRAE was 154  $\mu\text{m}$ , AVR was 0.679, and vessel density was 0.00689. We examined age- and sex-related differences in retinal and optic nerve head measures. Cup-to-disc ratio increased with age, whereas CRAE, AVR, and vessel density declined from early adulthood (Fig 3). Sex differences were statistically significant but small in magnitude: women showed slightly higher AVR and vessel density and lower cup-to-disc ratios at younger ages, with age-related differences largely similar between sexes and convergence most evident for cup-to-disc ratio at older ages (see Supplementary Tables 7–10, available at [www.ophtalmologyscience.org](http://www.ophtalmologyscience.org)).

## Deep Learning Models for Age, Sex, and Blood Pressure Prediction

To demonstrate the data set's potential for AI-based ophthalmology and oculomics research, we trained various deep learning models including convolutional and transformer architectures to predict age, sex, and blood pressure from the fundus images based on the data recorded during the baseline examination. These models included ResNet18, ResNet50, Inceptionv3, Xception, EfficientNetv2-Small, ViT-Small, and Swin-Tiny. We found that all architectures could predict these variables with similar performance compared to the literature<sup>24</sup> (Table 3 and Table 4). For instance, for age prediction, the MAE ranged from 2.964 years (2.920–3.005 95% CI) for the EfficientNet to 3.181 years (3.138–3.223) for the

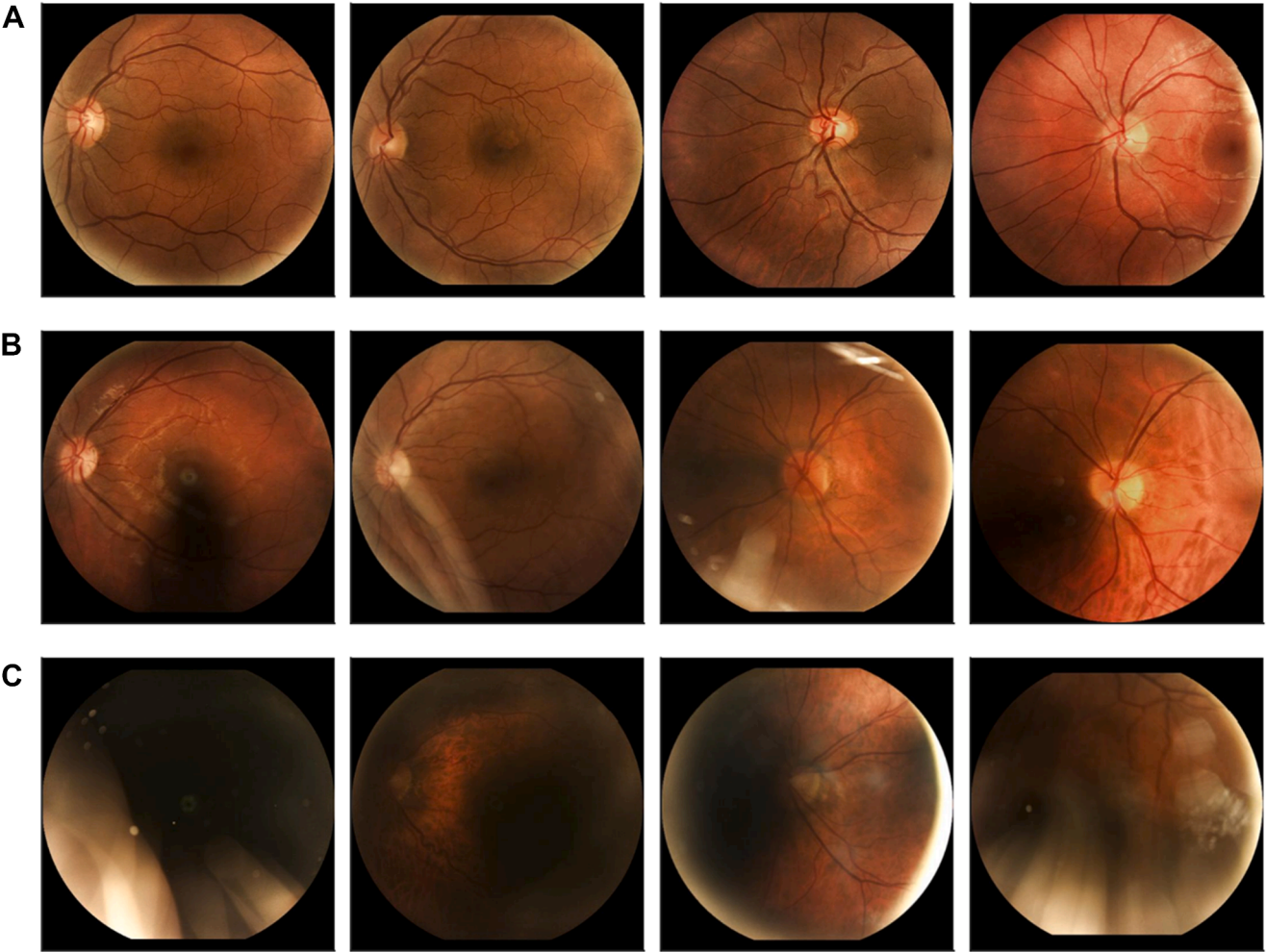


**Figure 2.** Example images from the NAKO data set. **A**, Fundus images with good quality, as judged by automatic quality grading by the Fundus Image Toolbox algorithm. From left to right: female, 47 years; male, 46 years; male, 32 years; and male, 27 years (as reported during the examination). **B**, Fundus images with borderline quality, showing evidence of artefacts or poor illumination. From left to right: female, 30 years; female, 46 years; female, 73 years; male, 48 years. **C**, Fundus images with insufficient quality not used for further analysis. From left to right: male, 53 years; male, 64 years; male, 63 years; and male, 54 years. NAKO = German National Cohort.

Xception network (Fig 4). For systolic blood pressure prediction, MAE ranged from 10.781 mmHg (10.644–10.922, 95% CI) for the EfficientNet to 11.419 mmHg (11.274–11.585, 95% CI) for the Swin

transformer, indicating substantial unexplained variability. The EffcientNetv2 network and the ResNet50 typically achieved the lowest error (Tables 3 and 4), making them good starting points for oculomics approaches.

Table 2. Percentage of Participants (n = 48 460) with Gradable Images According to 4 Different Artificial Intelligence Models for Quality Grading (FIT, RetiRatey, AutoMorph, and VascX) and Consensus Rating for Participants with Gradable Images According to all Models (All)

| Field         | FIT  | RetiRatey | AutoMorph | VascX | All  |
|---------------|------|-----------|-----------|-------|------|
| Right central | 83.1 | 90.4      | 85.4      | 87.7  | 81.3 |
| Right nasal   | 83.6 | 91.4      | 80.3      | 72.3  | 68.2 |
| Left central  | 76.7 | 84.7      | 77.9      | 82.3  | 74.9 |
| Left nasal    | 72.6 | 89.2      | 70.1      | 64.9  | 58.5 |

FIT = Fundus Image Toolbox.

## Discussion

We characterized the design and methodological framework of the ophthalmology module of the NAKO<sup>5</sup> and present baseline results from 48 460 participants who underwent standardized visual acuity testing and nonmydriatic color fundus imaging, as well as extensive biomedical examination. With its scale, NAKO is among the major international population-based cohorts in ophthalmology such as the UKB baseline ophthalmology subset with around 68 000 participants with OCT and color fundus imaging.<sup>23,25</sup> Within Germany, the study complements more local, well-established cohorts in Germany with

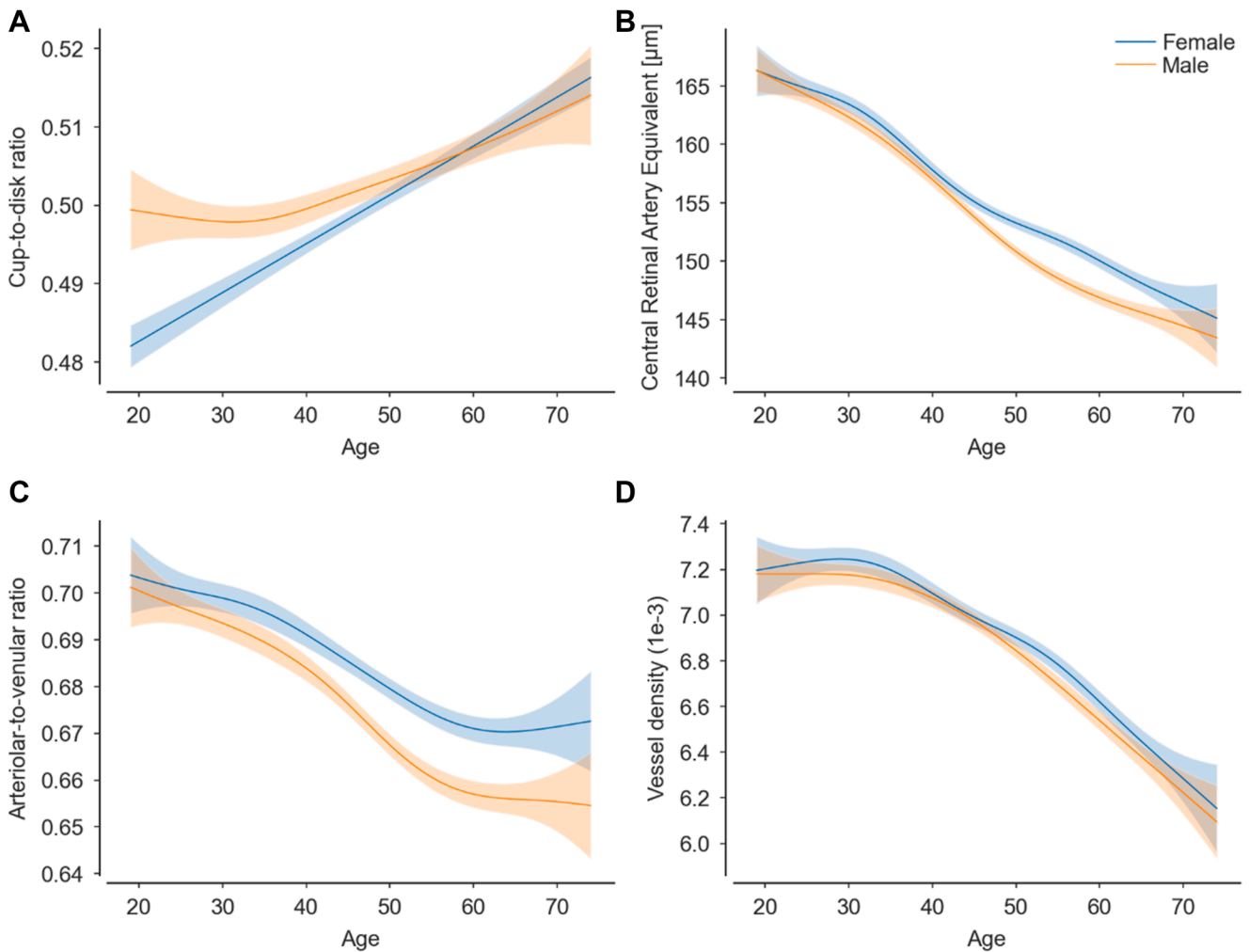


**Figure 3.** Morphological features extracted from fundus images show dependence on life course and sex. **A**, Cup-to-disk ratio. Solid lines show mean prediction of a fitted generalized additive model, with standard errors.  $n = 28\,220$ . **B**, Central retinal artery equivalent Knudtson.  $n = 27\,297$ . **C**, Arteriolar-to-venular ratio Knudtson.  $n = 27\,114$ . **D**, Vessel density.  $n = 28\,608$ . Detailed statistical information about the fitted generalized additive model models is available in the [Supplementary Material](#).

extensive phenotyping and ophthalmic assessments, such as the Gutenberg Health Study,<sup>26,27</sup> the AugUR study,<sup>28,29</sup> the Hamburg City Health Study,<sup>30,31</sup> the LIFE-Adult Study,<sup>32</sup> the SHIP study,<sup>33</sup> the Rhineland Study,<sup>34–36</sup> and the KORA study<sup>37,38</sup> (see [Table 5](#) for systematic comparison). In addition to the level 2 ophthalmology module, the NAKO includes local in-depth level 3 modules with ophthalmic relevance (such as for dry eye disease) in selected study centers.

The mean logarithm of the minimum angle of resolution values, which were close to 0.0, indicate that this community-based sample predominantly has normal vision. These values are broadly consistent with values reported in large European population-based cohorts, particularly among individuals without significant ocular pathology.<sup>23,26</sup> These visual acuity assessments in NAKO will become particularly important when follow-up investigations will be able to compare changes over time as individuals age and derive determinants of aging. Similarly, color fundus

images at baseline jointly with follow-up imaging will also be key to derive determinants of eye health changes as well as ocular presentations of other diseases. Together with objectively measured eye parameters, self-reported eye disease can provide insights into the awareness of eye diseases in the general population.

Self-reported eye disease can provide insight into disease awareness in the population, although it is likely to underestimate true disease burden, particularly for early or asymptomatic conditions, and should therefore be interpreted with caution.<sup>26</sup> As (semi-) automated image-based approaches evolve, they may complement basic self-reported disease assessment and resource and time intensive manual imaging grading for prevalence and incidence estimation by enabling more scalable phenotyping in especially larger data sets. Nonetheless, self-reported prevalence of cataract (4.2%) is consistent with prior evidence indicating that cataract prevalence is often in the single digits in younger populations and increases sharply

Table 3. Performance of Different Deep Learning Architectures for Prediction of Continuous Variables (Age, Systolic Blood Pressure, and Diastolic Blood Pressure), Measured by the MAE and  $R^2$ , with 95% Confidence Intervals

| Target                   | Model                      | MAE                    | $R^2$               |
|--------------------------|----------------------------|------------------------|---------------------|
| Age                      | Resnet18                   | 3.104 (3.063–3.148)    | 0.889 (0.885–0.893) |
|                          | Resnet50                   | 3.079 (3.037–3.121)    | 0.893 (0.889–0.896) |
|                          | Inceptionv3                | 3.113 (3.074–3.154)    | 0.889 (0.885–0.893) |
|                          | Xception                   | 3.181 (3.138–3.223)    | 0.885 (0.881–0.889) |
|                          | EfficientNetv2-S           | 2.964 (2.920–3.005)    | 0.901 (0.898–0.904) |
|                          | ViT-Small                  | 3.101 (3.057–3.144)    | 0.890 (0.886–0.894) |
|                          | Swin-Tiny                  | 3.079 (3.038–3.121)    | 0.894 (0.891–0.898) |
|                          | Poplin et al (Inceptionv3) | 3.26 (3.22, 3.31)      | 0.74 (0.73, 0.75)   |
| Systolic blood pressure  | Resnet18                   | 10.923 (10.789–11.052) | 0.210 (0.196–0.223) |
|                          | Resnet50                   | 10.853 (10.711–10.995) | 0.210 (0.196–0.224) |
|                          | Inceptionv3                | 10.950 (10.807–11.101) | 0.199 (0.186–0.213) |
|                          | Xception                   | 10.906 (10.764–11.049) | 0.188 (0.174–0.202) |
|                          | EfficientNetv2-S           | 10.781 (10.644–10.922) | 0.215 (0.201–0.229) |
|                          | ViT-Small                  | 10.970 (10.829–11.109) | 0.200 (0.187–0.212) |
|                          | Swin-Tiny                  | 11.419 (11.274–11.585) | 0.097 (0.083–0.111) |
|                          | Poplin et al (Inceptionv3) | 11.35 (11.18, 11.51)   | 0.36 (0.35, 0.37)   |
| Diastolic blood pressure | Resnet18                   | 6.969 (6.890–7.051)    | 0.204 (0.190–0.218) |
|                          | Resnet50                   | 6.970 (6.885–7.051)    | 0.205 (0.192–0.217) |
|                          | Inceptionv3                | 7.090 (7.006–7.171)    | 0.180 (0.167–0.193) |
|                          | Xception                   | 7.067 (6.976–7.152)    | 0.186 (0.171–0.199) |
|                          | EfficientNetv2-S           | 7.014 (6.931–7.099)    | 0.201 (0.187–0.214) |
|                          | ViT-Small                  | 6.967 (6.888–7.049)    | 0.204 (0.191–0.217) |
|                          | Swin-Tiny                  | 7.242 (7.155–7.327)    | 0.130 (0.117–0.143) |
|                          | Poplin et al (Inceptionv3) | 6.42 (6.33, 6.52)      | 0.32 (0.30, 0.33)   |

MAE = mean absolute error.

For comparison, literature values from a model trained by Poplin et al on the UK Biobank and EyePACS and evaluated on a holdout set from the UK Biobank.

with age, reaching high levels in older adults. Likewise, the prevalence of self-reported glaucoma (2.0%) aligns with population-based estimates reported in large European cohorts, where glaucoma prevalence generally ranges from approximately 1% to 3%, depending on age and diagnostic definition.<sup>39–41</sup> The prevalence of self-reported macular degeneration (0.9%) is low compared with image-based estimates in European cohorts, where age-related macular degeneration (AMD) prevalence increased with age from 3.5% to 17.6% for early AMD and from 0.1% to 9.8% for late AMD, likely because of the comparably young age of

the cohort and the occurrence of symptoms in intermediate and late AMD, while patients with early AMD have little to no symptoms.

Although image quality assessment remains a challenge due to the absence of internationally harmonized standards, we showed that automated quality-control tools can efficiently grade the large collections of fundus images in the NAKO data set. We found that the tools used in this study largely agree on image quality and that their ratings also agree well with expert ratings, despite detectable domain shifts between the training data sets of these algorithms and

Table 4. Performance of Different Deep Learning Architectures for Prediction of the Binary Variable Sex, Measured by the Accuracy and AUC, with 95% Confidence Intervals

| Target | Model                      | Accuracy            | AUC                 |
|--------|----------------------------|---------------------|---------------------|
| Sex    | Resnet18                   | 0.824 (0.818–0.830) | 0.908 (0.903–0.912) |
|        | Resnet50                   | 0.825 (0.818–0.831) | 0.910 (0.905–0.914) |
|        | Inceptionv3                | 0.829 (0.823–0.835) | 0.912 (0.907–0.917) |
|        | Xception                   | 0.825 (0.819–0.831) | 0.907 (0.903–0.912) |
|        | EfficientNetv2-S           | 0.835 (0.829–0.841) | 0.918 (0.914–0.922) |
|        | ViT-Small                  | 0.820 (0.814–0.826) | 0.903 (0.899–0.908) |
|        | Swin-Tiny                  | 0.836 (0.830–0.842) | 0.919 (0.915–0.923) |
|        | Poplin et al (Inceptionv3) | -                   | 0.97 (0.966, 0.971) |

AUC = area under the receiver operating characteristic curve.

For comparison, literature values from Poplin et al, trained on the UK Biobank and EyePACS and evaluated on a holdout set from the UK Biobank.

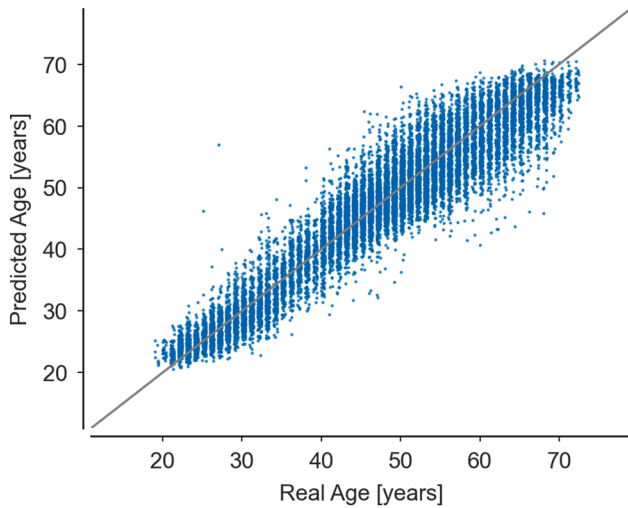


**Figure 4.** Real vs. predicted age for  $n = 15\,245$  NAKO participants for an EfficientNetv2-S model.  $r^2 = 0.901$ , MAE = 2.965. MAE = mean absolute error; NAKO = German National Cohort.

the NAKO data set. Remaining interalgorithm variability points to the need for further validation, transparency, and consensus guidance.

The large number of high-quality color fundus images from NAKO participants is a key strength of the data set and provides an opportunity to develop AI models for eye health research and ophthalmic approaches. However, commercial applications require separate disclosure and approval, and fundus images may only be used in publications in a manner that precludes reidentification of participants. The high quality of the fundus images also made it possible to use deep learning–based tools for automatically

extracting morphological features. This allowed us to analyze morphological differences of the retina at a larger scale to answer epidemiological questions, as illustrated for age and sex. Quality assessments by the AI models will be made available via the NAKO TransferHub to facilitate further research.

We further demonstrated that the NAKO data set is suited for developing deep learning techniques to study eye health, including ophthalmic approaches. Baseline deep learning models trained for predicting age, sex, and blood pressure performed similarly to the other approaches based on different fundus image data sets.<sup>24,42–44</sup> For instance, Poplin et al<sup>24</sup> trained an Inception-v3 model on subsets of the UKB and the EyePACS data set<sup>45</sup> to predict cardiovascular risk factors, including sex, age, and blood pressure. They obtained an area under the receiver operating characteristic curve of 0.97 for sex classification, MAE of 3.26 for age prediction, MAE of 11.35 for systolic blood pressure prediction, and MAE of 6.42 for diastolic blood pressure prediction, comparable to the aforementioned values we obtained. Interestingly, our experiments indicate that a ResNet18 (11.2 M), with a comparably small number of parameters, is sufficient to achieve state-of-the-art results, suggesting larger models could be overparametrized for these specific tasks.

The age-related increase in cup-to-disc ratio and the decline in retinal vessel calibers and AVR observed in this study are consistent with prior population-based and photographic studies describing structural remodeling of the optic nerve head and retina with aging.<sup>46,47</sup> Sex differences were limited and broadly consistent with previous reports.<sup>47</sup> These findings should be interpreted in light of the use of measurements of morphological properties obtained using the AutoMorph,<sup>11</sup> which, while scalable and reproducible,

**Table 5.** Comparison of the NAKO Ophthalmology Module with the UK Biobank and Regional German Population-Based Cohorts, Summarizing for Each Cohort Its Geographic Scope, Baseline Sample Size, Age Range, and the Ophthalmic Modalities Assessed (Visual Acuity, Refraction, Color Fundus Photography, OCT, and Additional Modalities)

| Cohort     | Scope    | Sample (Baseline)             | Age   | VA | Refraction | CFP | OCT | Additional Modalities | Reference                                  |
|------------|----------|-------------------------------|-------|----|------------|-----|-----|-----------------------|--------------------------------------------|
| NAKO       | National | ~49 000                       | 19–74 | X  |            | X   |     |                       | Present study                              |
| UK Biobank | National | ~67 000*                      | 40–69 | X  | X          | X   | X   | X                     | Chua 2019                                  |
| GHS        | Regional | ~15 000                       | 35–74 | X  | X          | X   | X   | X                     | Höhn 2015                                  |
| HCHS       | Regional | ~10 000 (45 000) <sup>†</sup> | 45–74 | X  | X          | X   | X   | X                     | Beuse 2025                                 |
| LIFE-Adult | Regional | ~10 000                       | 40–79 |    |            | X   | X   |                       | Loeffler 2015                              |
| Rhineland  | Regional | ~5000 (20 000) <sup>‡</sup>   | 30–95 | X  | X          | X   | X   | X                     | Garzone 2023, Garzone 2025, Mauschitz 2019 |
| SHIP       | Regional | ~4000                         | 20–79 |    |            | X   |     |                       | Jürgens 2014                               |
| KORA       | Regional | ~3000                         | 25–74 |    |            | X   |     |                       | Brandl 2016                                |
| AugUR      | Regional | ~3000                         | ≥70   | X  |            | X   | X   | X                     | Stark 2015, Brandl 2018                    |

AugUR = age-related diseases: understanding genetic and non-genetic influences – a study at the University of Regensburg; CFP = color fundus photography; GHS = Gutenberg Health Study; HCHS = Hamburg City Health Study; KORA = Cooperative Health Research in the Region of Augsburg; LIFE = Leipzig Research Center for Civilization Diseases; NAKO = German National Cohort; SHIP = Study of Health in Pomerania; UK = United Kingdom; VA = visual acuity.

\*A total of ~117 000 participants were screened for vision, intraocular pressure, refraction and keratometry, while a subgroup of ~67 000 underwent eye imaging.

<sup>†</sup>Ongoing: published analyses based of the first ~10 000 of a planned 45 000 cohort.

<sup>‡</sup>Ongoing: published analyses based of the first ~5000 of a planned 20 000 cohort.

require further validation against manual expert segmentation, as manual annotation remains the current reference standard. Also, the lack of correction for axial length differences could affect these results (see below). Automatically measured morphological properties will be made available via the NAKO TransferHub as derived variables from the data.

Beyond the baseline characterization presented here, the ophthalmic data available within NAKO will continue to grow, as data from follow-up examinations will become available. In addition, manual image gradings and imaging-derived phenotypes will be released, thereby expanding the scientific scope of the data. Follow-up analyses will permit the assessment of changes over time, and previously developed and adapted AI-based methods will enable the efficient provision of further derived variables for use in scientific analyses and for identifying determinants of progression of ocular disease or maintenance of ocular health at older age.

## Strengths and Limitations

Key strengths of NAKO are its large scale, nationwide coverage, and longitudinal design. Standardized operating procedures across study centers, together with high-quality nonmydriatic color fundus photography and visual acuity assessment, support robust population-based analyses. These data enable the investigation of spatial differences and temporal changes. While this present study focused on automated image analysis pipelines, future releases of manually graded imaging data will improve disease-specific eye phenotyping and validation of imaging-derived endpoints. Recent advances in ophthalmics and AI highlight the potential of retinal imaging as a biomarker for ocular and systemic diseases. However, its large-scale application in population-based cohorts remains limited. The extensive phenotyping in NAKO will allow for detecting unique, new associations of color fundus image features with a broad set of diseases and health conditions.

Several limitations should be acknowledged. First, retinal imaging within the NAKO ophthalmic assessment is limited to nonmydriatic color fundus photography without complementary modalities, such as OCT or widefield imaging. This restricts detailed structural in depth and peripheral phenotyping. Reliance on self-reported ocular disease may also result in misclassification and underascertainment.<sup>48,49</sup> Second, we could not correct our morphological measurements for magnification differences due to axial length, refractive error, or corneal curvature, because this information was not available in the data set. Ratio-based measurements like AVR are inherently robust

to this, but absolute measurements like CRAE, as well as the reported sex-related differences for these measurements, should be interpreted with care. Third, the width of non-imaging ophthalmic assessment was relatively narrow, encompassing only visual acuity. Fourth, ethnicity data are not available for analysis for NAKO data sets, limiting analyses into fairness and bias in this regard.

## Conclusion and Outlook

The ophthalmic assessment within the NAKO provides a large, standardized, prospective, population-based resource integrating retinal imaging, visual acuity, and comprehensive systemic health measures. We demonstrated its potential as a modern, accessible platform for methodological development and validation in eye health research. Future analyses including follow-up data will provide the opportunity to investigate longitudinal developments among NAKO participants and use the results of the ophthalmic assessment in NAKO for early detection of ocular or systemic diseases.

## Declaration of Generative AI and AI-Assisted Technologies in the Writing Process:

During the preparation of this work the author(s) used DeepL Write and Claude Opus 4.8 in order to improve style and grammar. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

## Acknowledgments

The authors thank Daniel Zeymer von Metnitz and Immanuel Seitz for help with quality grading. This project was conducted with data from the German National Cohort (NAKO; [www.nako.de](http://www.nako.de); Application No. NAKO-590 and NAKO-810). The NAKO is funded by the Federal Ministry of Research, Technology and Space (BMFTR; project funding reference numbers: 01ER1301A/B/C, 01ER1511D, 01ER1801A/B/C/D, and 01ER2301A/B/C), federal states of Germany and the Helmholtz Association, the participating universities and the institutes of the Leibniz Association. The authors thank all participants who took part in the NAKO study and the staff of this research initiative. PB is a member of the Excellence Cluster 2064 “Machine Learning – New Perspectives for Science” and the NAKO AI expert group. AS, MU, AKS, CB, IH, GG, RF are members of the NAKO retina expert group. AB is a member of the Digital Clinician Scientist Program at University Medical Center Hamburg-Eppendorf (UKE).

## Footnotes and Disclosures

Originally received: May 7, 2026.

Final revision: September 1, 2026.

Accepted: September 1, 2026.

Available online: September 11, 2026. Manuscript no. XOPS-D-26-00545.

<sup>1</sup> Hertie Institute for AI in Brain Health, University of Tübingen, Tübingen, Germany.

<sup>2</sup> Tübingen AI Center, University of Tübingen, Tübingen, Germany.

<sup>3</sup> Department of Ophthalmology, University Medical Center Hamburg-Eppendorf (UKE), Hamburg, Germany.

<sup>4</sup> University Eye Hospital, Centre for Ophthalmology, University of Tübingen, Tübingen, Germany.

<sup>5</sup> Institute of Epidemiology and Social Medicine, University of Münster, Münster, Germany.

<sup>6</sup> Department of Genetic Epidemiology, University of Regensburg, Regensburg, Germany.

<sup>7</sup> Department of Ophthalmology, University of Regensburg, Regensburg, Germany.

<sup>8</sup> German Cancer Research Center (DKFZ) Heidelberg, Division of Cancer Epidemiology, Heidelberg, Germany.

<sup>9</sup> Department of Epidemiological Methods and Etiological Research, Leibniz Institute for Prevention Research and Epidemiology — BIPS, Bremen, Germany.

<sup>10</sup> Department of Ophthalmology, Medical Faculty Mannheim, University of Heidelberg, Mannheim, Germany.

<sup>11</sup> Department of Ophthalmology, Heinrich-Heine University, Düsseldorf, Germany.

<sup>12</sup> Institute of Social Medicine, Epidemiology and Health Economics, Charité - Universitätsmedizin Berlin, Berlin, Germany.

<sup>13</sup> Institute of Clinical Epidemiology and Biometry, University of Würzburg, Würzburg, Germany.

<sup>14</sup> State Institute of Health I, Bavarian Health and Food Safety Authority, Erlangen, Germany.

<sup>15</sup> Helmholtz Center for Infection Research, Braunschweig, Germany.

<sup>16</sup> Institute for Epidemiology and Prevention Medicine, University of Regensburg, Regensburg, Germany.

<sup>17</sup> Institut für Community Medicine, Universitätsmedizin Greifswald, Greifswald, Germany.

<sup>18</sup> Chair of Epidemiology, IBE, Faculty of Medicine, LMU Munich, Germany.

<sup>19</sup> Helmholtz Zentrum München, German Research Center for Environmental Health (GmbH), Institute of Epidemiology, Neuherberg, Germany.

<sup>20</sup> Department of Molecular Epidemiology, German Institute of Human Nutrition Potsdam-Rehbrücke, Nuthetal, Germany.

<sup>21</sup> Institute of Nutritional Science, University of Potsdam, Nuthetal, Germany.

<sup>22</sup> Department of Ophthalmology, University Medical Center of the Johannes Gutenberg-University Mainz, Mainz, Germany.

<sup>23</sup> Institute for Ophthalmic Research, Centre for Ophthalmology, University of Tübingen, Tübingen, Germany.

\*C.R. and A.B. are co-first authors.

\*\*M.U. and P.B. are co-senior authors.

Disclosure(s):

All authors have completed and submitted the ICMJE disclosures form.

The author(s) have made the following disclosure(s):

A.S.: Grants — University Eye Hospital Tuebingen; Consultant — Bayer, Novartis, Roche; Honoraria — Bayer, Heidelberg Engineering, Roche; Travel expenses — Bayer, Roche.

R.F.: Consultant — Alimera, Apellis, Astellas, Bayer, Biogen, Böhlinger-Ingelheim, Catena, Novartis, ODOS, ProGenerika, Stada Pharm, Roche; Payment for expert testimony — Roche.

P.B.: Grants — European Research Council, German Ministry for Research, Technology and Space (BMFTR); Travel expenses — State of Hesse, Ministry for Science.

The project was funded by the Hertie Foundation and DFG.

**HUMAN SUBJECTS:** Human subjects were included in this study. The study protocol of NAKO was approved by the ethics committees of all participating institutions, and all participants provided written informed consent in accordance with German legal and data protection requirements. It is conducted in accordance with the Declaration of Helsinki and national standards for good clinical and epidemiological practice (Peters et al., 2022).

No animal subjects were used in this study.

**Author Contributions:**

Conception and design: Roa, Beuse, Schweig, Berger, Brandl, Brinker, Elbrecht, Finger, Geerling, Grohmann, Greiser, Günther, Heid, Karch, Keil, Krepel, Leitzmann, Meinke-Franze, Peters, Schipf, Schulz, Schuster, Willich, Leitritz, Ueffing, Berens

Analysis and interpretation: Roa, Beuse, Müller, Berens

Data collection: Roa, Beuse, Schweig, Berger, Brandl, Brinker, Elbrecht, Finger, Geerling, Grohmann, Greiser, Günther, Heid, Karch, Keil, Krepel, Leitzmann, Meinke-Franze, Peters, Schipf, Schulz, Schuster, Willich, Leitritz, Ueffing, Berens

Obtained funding: Schweig, Berger, Greiser, Heid, Karch, Keil, Leitzmann, Peters, Schipf, Schulz, Willich, Ueffing, Berens, Beuse

Overall responsibility: Roa, Beuse, Schweig, Berens

**Abbreviations and Acronyms:**

**AI** = artificial intelligence; **AMD** = age-related macular degeneration; **AVR** = arteriolar-to-venular ratio; **CI** = confidence interval; **CRAE** = central retinal artery equivalent; **FIT** = Fundus Image Toolbox; **IQR** = interquartile range; **MAE** = mean absolute error; **NAKO** = German National Cohort; **UKB** = UK Biobank.

**Keywords:**

Artificial intelligence, Epidemiology, Prospective cohort study, Fundus image, Quality control.

**Correspondence:**

Philipp Berens, PhD, Hertie Institute for AI in Brain Health, University of Tübingen, Maria-von-Linden-Str. 6, 72076 Tübingen, BW, Germany. E-mail: [philipp.berens@uni-tuebingen.de](mailto:philipp.berens@uni-tuebingen.de).

## References

1. Peters A, Peters A, Greiser KH, et al. Framework and baseline examination of the German National Cohort (NAKO). *Eur J Epidemiol.* 2022;37(10):1107–1124.
2. Schipf S, Schöne G, Schmidt B, et al. Die Basiserhebung der NAKO Gesundheitsstudie: teilnahme an den Untersuchungsmodulen, Qualitätssicherung und Nutzung von Sekundärdaten. *Bundesgesundheitsblatt - Gesundheitsforschung - Gesundheitsschutz.* 2020;63(3):254–266.
3. Sudlow C, Gallacher J, Allen N, et al. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old Age. *PLOS Med.* 2015;12(3):e1001779.
4. Chen Z, Chen J, Collins R, et al. China Kadoorie Biobank of 0.5 million people: survey methods, baseline characteristics and long-term follow-up. *Int J Epidemiol.* 2011;40(6):1652–1666.
5. The National German Cohort Consortium. The German National Cohort: aims, study design and organization. *Eur J Epidemiol.* 2014;29(5):371–382.
6. Rach S, Sand M, Reineke A, et al. The baseline examinations of the German National Cohort (NAKO): Recruitment protocol, response, and weighting. *Eur J Epidemiol.* 2025;40(4):475–489.
7. Leitritz MA, Hense H-W, Schiefer U, et al. [Development and first results of fast and cost-effective examination methods for

- an ophthalmological screening within the National Cohort]. *Klinische Monatsblätter für Augenheilkunde*. 2013;230(12):1238–1246.
8. Gervelmeyer J, Müller S, Huang Z, Berens P. Fundus Image Toolbox: a Python package for fundus image processing. *J Open Source Softw*. 2025;10(108):7101.
9. Liu R, Wang X, Wu Q, et al. DeepDRiD: diabetic Retinopathy—Grading and image quality estimation challenge. *Patterns*. 2022;3(6):100512.
10. Oraá RR, García M, Oraá-Pérez J, et al. Automatic fundus image quality assessment: diagnostic accuracy in clinical practice. *Invest Ophthalmol Vis Sci*. 2020;61(7):2033.
11. Zhou Y, Wagner SK, Chia MA, et al. AutoMorph: automated retinal vascular morphology quantification via a deep learning pipeline. *Transl Vis Sci Technol*. 2022;11(7):12.
12. Fu H, Wang B, Shen J, et al. Evaluation of retinal image quality assessment networks in different color-spaces. In: Shen D, Liu T, Peters TM, et al., eds. *Medical Image Computing and Computer Assisted Intervention — MICCAI 2019*. Cham, Switzerland: Springer International Publishing; 2019:48–56.
13. Vargas Quiros J, Liefers B, van Garderen KA, et al. VascX models: deep ensembles for retinal vascular analysis from color fundus images. *Transl Vis Sci Technol*. 2025;14(7):19.
14. Knudtson MD, Lee KE, Hubbard LD, et al. Revised formulas for summarizing retinal vessel diameters. *Curr Eye Res*. 2003;27(3):143–149.
15. Wood SN. *Generalized Additive Models: An Introduction with R, Second Edition*. 2nd ed. New York: Chapman and Hall/CRC; 2017. <https://doi.org/10.1201/9781315370279>.
16. Chollet F. *Xception: Deep Learning With Depthwise Separable Convolutions*; 2017:1251–1258. [https://openaccess.thecvf.com/content\\_cvpr\\_2017/html/Chollet\\_Xception\\_Deep\\_Learning\\_CVPR\\_2017\\_paper.html](https://openaccess.thecvf.com/content_cvpr_2017/html/Chollet_Xception_Deep_Learning_CVPR_2017_paper.html). Accessed October 24, 2025.
17. Dosovitskiy A, Beyer L, Kolesnikov A, et al. An image is worth 16x16 words: transformers for image recognition at Scale. International Conference on Learning Representations. <https://openreview.net/forum?id=YicbFdNTTy>; 2020. Accessed October 24, 2025.
18. He K, Zhang X, Ren S, Sun J. *Deep Residual Learning for Image Recognition*; 2016:770–778. [https://openaccess.thecvf.com/content\\_cvpr\\_2016/html/He\\_Deep\\_Residual\\_Learning\\_CVPR\\_2016\\_paper.html](https://openaccess.thecvf.com/content_cvpr_2016/html/He_Deep_Residual_Learning_CVPR_2016_paper.html). Accessed October 24, 2025.
19. Liu Z, Lin Y, Cao Y, et al. *Swin transformer: hierarchical vision transformer using shifted windows*; 2021:10012–10022. [https://openaccess.thecvf.com/content\\_ICCV2021/html/Liu\\_Swin\\_Transformer\\_Hierarchical\\_Vision\\_Transformer\\_-\\_Using\\_Shifted\\_Windows\\_ICCV\\_2021\\_paper.html](https://openaccess.thecvf.com/content_ICCV2021/html/Liu_Swin_Transformer_Hierarchical_Vision_Transformer_-_Using_Shifted_Windows_ICCV_2021_paper.html). Accessed October 24, 2025.
20. Szegedy C, Vanhoucke V, Ioffe S, et al. *Rethinking the Inception Architecture for Computer Vision*; 2016:2818–2826. [https://www.cv-foundation.org/openaccess/content\\_cvpr\\_2016/html/Szegedy\\_Rethinking\\_the\\_Inception\\_CVPR\\_2016\\_paper.html](https://www.cv-foundation.org/openaccess/content_cvpr_2016/html/Szegedy_Rethinking_the_Inception_CVPR_2016_paper.html). Accessed October 24, 2025.
21. Tan M, Le Q. EfficientNetV2: smaller models and faster training. In: *Proceedings of the 38th International Conference on Machine Learning*; 2021:10096–10106. <https://proceedings.mlr.press/v139/tan21a.html>. Accessed October 24, 2025.
22. Stein MJ, Fischer B, Bohmann P, et al. Differences in anthropometric measures based on sex, Age, and health status. *Deutsches Ärzteblatt Int*. 2024;121(7):207–213.
23. Chua SYL, Thomas D, Allen N, et al. Cohort profile: design and methods in the eye and vision consortium of UK Biobank. *BMJ Open*. 2019;9(2):e025077.
24. Poplin R, Varadarajan AV, Blumer K, et al. Prediction of cardiovascular risk factors from retinal fundus photographs via deep learning. *Nat Biomed Eng*. 2018;2(3):Article.3.
25. Warwick AN, Curran K, Hamill B, et al. UK Biobank retinal imaging grading: methodology, baseline characteristics and findings for common ocular diseases. *Eye*. 2023;37(10):2109–2116.
26. Höhn R, Kottler U, Peto T, et al. The ophthalmic branch of the Gutenberg Health Study: study design, cohort profile and self-reported diseases. *PLOS ONE*. 2015;10(3):e0120476.
27. Wild PS, Zeller T, Beutel M, et al. [The Gutenberg Health Study]. *Bundesgesundheitsblatt, Gesundheitsforschung, Gesundheitsschutz*. 2012;55(6–7):824–829.
28. Brandl C, Brücklmayer C, Günther F, et al. Retinal layer thicknesses in early age-related macular degeneration: results from the German AugUR Study. *Invest Ophthalmol Vis Sci*. 2019;60(5):1581–1594.
29. Stark K, Olden M, Brandl C, et al. The German AugUR study: study protocol of a prospective study to investigate chronic diseases in the elderly. *BMC Geriatr*. 2015;15(1):130.
30. Beuse A, Hoppert I, Wolfram C, et al. Hamburg City Health Study: design and ocular baseline characteristics of the first 10,000 participants. *Invest Ophthalmol Vis Sci*. 2025;66(8), 2009–2009.
31. Jagodzinski A, Johansen C, Koch-Gromus U, et al. Rationale and design of the Hamburg City health Study. *Eur J Epidemiol*. 2020;35(2):169–181.
32. Loeffler M, Engel C, Ahnert P, et al. The LIFE-Adult-Study: objectives and design of a population-based cohort study with 10,000 deeply phenotyped adults in Germany. *BMC Public Health*. 2015;15(1):691.
33. Jürgens C, Völzke H, Tost F. Study of health in Pomerania (SHIP-Trend). *Der Ophthalmologe*. 2014;111(5):443–447.
34. Garzone D, Finger RP, Mauschitz MM, et al. Visual impairment and retinal and brain neurodegeneration: a population-based study. *Hum Brain Mapp*. 2023;44(7):2701–2711.
35. Garzone D, Imtiaz MA, Mauschitz MM, et al. Age-Related macular degeneration and its genetic risk: a population-based Study. *Curr Eye Res*. 2025;50(1):82–86.
36. Mauschitz MM, Lohner V, Koch A, et al. Retinal layer assessments as potential biomarkers for brain atrophy in the Rhineland Study. *Sci. Rep*. 2022;12(1):2757.
37. Brandl C, Breinlich V, Stark KJ, et al. Features of age-related macular degeneration in the general adults and their dependency on Age, sex, and smoking: results from the German KORA Study. *PLOS ONE*. 2016;11(11):e0167181.
38. Holle R, Happich M, Löwel H, Wichmann HE; Group, for the M. S. KORA - a research platform for population based health research. *Das Gesundheitswesen*. 2005;67(S 01):19–25.
39. Dielemans I, Vingerling JR, Wolfs RCW, et al. The prevalence of primary open-angle glaucoma in a population-based Study in the Netherlands: the rotterdam Study. *Ophthalmology*. 1994;101(11):1851–1855.
40. Shweikh Y, Ko F, Chan MPY, et al. Measures of socioeconomic status and self-reported glaucoma in the UK Biobank cohort. *Eye*. 2015;29(10):1360–1367.
41. Wolfs RCW, Borger PH, Ramrattan RS, et al. Changing views on open-angle glaucoma: definitions and Prevalences—The

- rotterdam Study. *Invest Ophthalmol Vis Sci.* 2000;41(11): 3309–3321.
42. Ahadi S, Wilson KA, Babenko B, et al. Longitudinal fundus imaging and its genome-wide association analysis provide evidence for a human retinal aging clock. *eLife.* 2023;12: e82364.
43. Yu Z, Chen R, Gui P, et al. A cross population study of retinal aging biomarkers with longitudinal pre-training and label distribution learning. *Npj Digital Med.* 2025;8(1):344.
44. Zhu Z, Wang Y, Qi Z, et al. Oculomics: current concepts and evidence. *Prog Retin Eye Res.* 2025;106:101350.
45. Cuadros J, Bresnick G. EyePACS: an adaptable telemedicine System for diabetic retinopathy screening. *J Diabetes Sci Technol.* 2009;3(3):509–516.
46. Garway-Heath DF, Wollstein G, Hitchings RA. Aging changes of the optic nerve head in relation to open angle glaucoma. *Br J Ophthalmol.* 1997;81(10):840–845.
47. Leung H, Wang JJ, Rochtchina E, et al. Relationships between Age, blood pressure, and retinal vessel diameters in an older population. *Invest Ophthalmol Vis Sci.* 2003;44(7): 2900–2904.
48. Foreman J, Xie J, Keel S, et al. The validity of self-report of eye diseases in participants with vision loss in the National Eye Health Survey. *Sci Rep.* 2017;7(1):8757.
49. Patty L, Wu C, Torres M, et al. Validity of self-reported eye disease and treatment in a population-based Study: the Los Angeles Latino eye Study. *Ophthalmology.* 2012;119(9): 1725–1730.