Genetic Risk Score Analysis Supports a Joint View of Two Classification Systems for Age-Related Macular Degeneration

Janina M. Herold,[1] Martina E. Zimmermann,[1] Mathias Gorski,[1] Felix Günther,[2] Bernhard H. F. Weber,[3,4] Horst Helbig,[5] Klaus J. Stark,[1] Iris M. Heid,[1] and Caroline Brandl[1,5]

[1]Department of Genetic Epidemiology, University of Regensburg, Franz-Josef-Strauß-Allee 11, Regensburg, Germany
[2]Department of Mathematics, Stockholm University, Albanovägen 28, Stockholm, Sweden
[3]Institute of Human Genetics, University of Regensburg, Franz-Josef-Strauß-Allee 11, Regensburg, Germany
[4]Institute of Clinical Human Genetics, University Hospital Regensburg, Franz-Josef-Strauß-Allee 11, Regensburg, Germany
[5]Department of Ophthalmology, University Hospital Regensburg, Franz-Josef-Strauß-Allee 11, Regensburg, Germany

Correspondence: Iris M. Heid, Department of Genetic Epidemiology, University of Regensburg, Franz-Josef-Strauss-Allee 11, Regensburg 93053, Germany; iris.heid@ukr.de.

IMH and CB supervised these analyses jointly.

Received: June 23, 2023
Accepted: August 25, 2023
Published: September 18, 2023

Citation: Herold JM, Zimmermann ME, Gorski M, et al. Genetic risk score analysis supports a joint view of two classification systems for age-related macular degeneration. *Invest Ophthalmol Vis Sci.* 2023;64(12):31. https://doi.org/10.1167/iovs.64.12.31

Purpose. The purpose of this study was to evaluate the utility of combining the Clinical Classification (CC) and the Three Continent age-related macular degeneration (AMD) Consortium Severity Scale (3CACSS) for classification of AMD.

Methods. In two independent cross-sectional datasets of our population-based AugUR study (Altersbezogene Untersuchungen zur Gesundheit der Universität Regensburg), we graded AMD via color fundus images applying two established classification systems (CC and 3CACSS). We calculated the genetic risk score (GRS) across 50 previously identified variants for late AMD, its association via logistic regression, and area under the curve (AUC) for each AMD stage.

Results. We analyzed 2188 persons aged 70 to 95 years. When comparing the two classification systems, we found a distinct pattern: CC "age-related changes" and CC "early AMD" distinguished individuals with 3CACSS "no AMD"; 3CACSS "mild/moderate/severe early AMD" stages, and distinguished CC "intermediate AMD". This suggested a 7-step scale combining the 2 systems: (i) "no AMD", (ii) "age-related changes", (iii) "very early AMD", (i.e. CC "early"), (iv) "mild early AMD", (v) "moderate early AMD", (vi) "severe early AMD", and (vii) "late AMD". GRS association and diagnostic accuracy increased stepwise by increased AMD severity in the 7-step scale and by increased restriction of controls (e.g. for CC "no AMD without age-related changes": AUC = 55.1%, 95% confidence interval [CI] = 51.6, 58.6, AUC = 62.3%, 95% CI = 59.1, 65.6, AUC = 63.8%, 95% CI = 59.3, 68.3, AUC = 78.1%, 95% CI = 73.6, 82.5, AUC = 82.2%, 95% CI = 78.4, 86.0, and AUC = 79.2%, 95% CI = 75.4, 83.0). A stepwise increase was also observed by increased drusen size and area.

Conclusions. The utility of a 7-step scale is supported by our clinical and GRS data. This harmonization and full data integration provides an immediate simplification over using either CC or 3CACSS and helps to sharpen the control group.

Keywords: age-related macular degeneration (AMD), population-based study, genetic risk score (GRS), drusen, classification systems, clinical classification (CC), three continent AMD consortium severity scale (3CACSS)

Late stage age-related macular degeneration (AMD) leads to severe vision impairment with limited options of therapeutic or preventive intervention.[1] Distinguishing risk factors for the development of early disease stages from factors for progression to late AMD are pivotal to help tailor the proper clinical strategies for patient management at each of the various stages of disease progression.

Currently, the investigation of particularly early AMD stages is hampered in epidemiological studies by different ways to define "early AMD".[2–5] Whereas the definition of late AMD as geographic atrophy (GA) and/or macular neovascular (MNV) complications[2–4] is reasonably homogeneous between classification systems, the "early AMD" definitions differ in how the systems combine AMD features like differently sized yellowish accumulations of extracellular material (drusen) or abnormalities of the retinal pigment epithelium (RPE; depigmentation or increased amount of pigment). Two frequently applied classification systems in observational studies are, for example, the Clinical Classification (CC)[4] and the Three Continent AMD Consortium Severity Scale (3CACSS).[2] The CC differentiates (i) no AMD without age-related changes, (ii) no AMD with age-related changes, (iii) "early AMD", and (iv) "intermediate AMD". The 3CACSS differentiates (i) no AMD, (ii) "mild early", (iii) "moderate

early". and (iv) "severe early AMD" (Supplementary Table S1). These differences render cross-study comparisons difficult and the terminology of "early" versus "intermediate" AMD is often giving rise to confusion. Previous work has explored the performance of several AMD classification systems and showed that each system has its merits in different settings.[6] A unified system has often been recognized as an important, but still unsolved challenge.[2]

We have previously observed a distinct pattern when comparing CC and 3CACSS in the cross-sectional data of the first part of our Altersbezogene Untersuchungen zur Gesundheit der Universität Regensburg, AugUR1 baseline survey (AugUR study), a population-based study in individuals aged ≥ 70 years[3] (see Supplementary Table S1): CC-based "age-related changes" and CC "early AMD" were fully distinct from 3CACSS "early AMD"; they distinguished stages within 3CACSS "no AMD". Moreover, 3CACSS-based "mild/moderate/severe early AMD" stages were distinct from CC "early AMD"; they distinguished stages within the CC "intermediate AMD". The CC system thus appeared to ignore relevant information when collapsing 3CACSS "early stages" into CC "intermediate AMD" – given the 3CACSS "early" stages have been shown to be associated with a step-wise increasing risk for late AMD.[6,7] The problem further extends to the control group definition: the 3CACSS system ignores information by collapsing CC "early AMD" and CC "age-related changes" into the 3CACSS "no AMD" group – and thus into the 3CACSS control group. Our previous data showed that "analyzing early AMD individuals" or "excluding early AMD individuals" meant something very different depending on the classification system.

To overcome the confusing terminology and the disregard of relevant information, we now suggest that these 2 classification systems can be combined in a meaningful way into a 7-step scale (see Supplementary Table S1): (i) CC "no AMD without age-related changes", (ii) CC "no AMD with age-related changes", (iii) CC "early AMD", (iv) 3CACSS "mild early", (v) 3CACSS "moderate early", and (vi) 3CACSS "severe early", and (vii) late AMD. In order to document the utility of this 7-step scale, we here aimed (1) to provide a replication for the previously observed distinct pattern when comparing the CC and 3CACSS classifications and (2) to provide a rationale of the stepwise gain in information with regard to risk prediction of late AMD.

A typical approach to document a stepwise gain of predicting the progression from early to late AMD is the quantification of the predictive ability. This is challenging as it requires large sample sizes with progression data on individuals with early AMD. Although the stepwise gain to predict late AMD has previously been shown for 3CACSS "mild/moderate/severe early AMD",[6,7] the documentation of at least some predictive ability of CC "early AMD" might have failed due to limited power. By making full use of the cross-sectional data and the strong genetic component of late AMD,[8] we additionally integrated the genetic risk score (GRS) based on 50 previously identified genetic variants for late AMD.[8] We evaluated whether the GRS showed a stepwise increased risk for each of the "early AMD" stages and whether different control group definitions sharpened these risk estimates. We thus (1) analyzed independent cross-sectional data using the second part of our AugUR study (AugUR2 baseline survey), to replicate the distinct pattern previously observed in AugUR1, and (2) computed the GRS association for case groups of increased AMD severity and control groups with increased restrictions in the approximately 2200 individuals of AugUR1 and 2.

Subjects and Methods

Study Population, Study Sample, and Data Collection

AugUR is designed as a population-based cohort study with the follow-up ongoing until 2025. Participants were recruited from the mobile elderly population in/around Regensburg, Germany, a study region of approximately 350,000 inhabitants of mostly Caucasian ancestry. AugUR recruitment was conducted in two independent baseline surveys (AugUR1, 2013-2015 and AugUR2, 2017-2019). Study recruitment and study conduct was the same for both surveys. No participants of AugUR1 were enrolled in AugUR2. Here, we present cross-sectional results from combining both baseline surveys.

Study recruitment and conduct were described previously.[3,7,9,10] Briefly, inhabitants of the city and county of Regensburg, Germany, with ≥ 70 years of age, were identified by local registries and invited by a mailed written invitation letter to the study center at the University Hospital Regensburg. Individuals invited to AugUR1 were excluded from the invitation list for AugUR2. Individuals were included as participants, if they were physically able and willing to come to the study center, to provide informed written consent and to participate in a 3-hour study program.

Details on data collection have been described elsewhere.[3,7,9,10] In brief, information on lifestyle factors, metabolic parameters, and general and ocular comorbidities were gathered via a standardized face-to-face interview, medical examinations by trained medical staff, and laboratory measurements from blood or urine.[3,7,9,10]

The AugUR study was approved by the Ethics Committee of the University of Regensburg, Germany (vote 12-101-0258). The study complies with the Declaration of Helsinki and its later amendments. All participants provided informed written consent.

Assessment of AMD-Related Features on Color Fundus Images and AMD Classification

Color fundus photography of the central retina and assessment of AMD features were conducted as described previously[3,7,10]; procedures and graders were the same for both surveys, AugUR1 and AugUR2. For each eye, the presence, size (small, intermediate, and large) and area of drusen, pigment abnormalities (hyperpigmentation or depigmentation), GA, including central or paracentral location, and/or MNV were determined using gradable color fundus images.

This information on AMD features was then transferred into the AMD stage per eye according to 2 established classification systems, the 3CACSS[2] and CC[4], as "no" (absence of "early" or "late"), "early" (3CACSS = mild, moderate, or severe early and CC = early or intermediate) or "late" AMD (details described previously[3]; see Supplementary Table S1). As defined by these two classification systems, only lesions within 2 standard disc diameters (approximately 3000 µm) of the center of the macula/fovea were considered; additionally, drusen outside this central area were documented as peripheral drusen. Although not part of the two classification systems, subretinal drusenoid deposits (SDDs) and

non-AMD-related pigmentary abnormalities were also assessed if present (see Supplementary Table S1).

Finally, the AMD status of a person was derived as the AMD status of the eye with the more severe stage ("worse eye") when color fundus images of both eyes were gradable for AMD. When images were gradable only for one eye, the AMD status of the person was the AMD stage of this eye. We analyzed individuals with at least one gradable eye.

Additionally, the AMD features per eye were also transferred into features per person by combining the AMD features from both eyes. Participants with gradable color fundus images from only one eye were excluded from this analysis. Different combinations of AMD features per person were defined as described as follows: (i) participants with no AMD features for both eyes were used as controls; (ii) single drusen types as small, peripheral drusen, or SDDs only; (iii) combined drusen types for intermediate or large drusen whereby intermediate drusen included presence of small drusen, large drusen included small, and intermediate drusen, in addition large drusen were separated by size based on the O2 circle; (iv) separated combinations of non-AMD- and AMD-related pigment abnormalities with different drusen types; and (v) late AMD separated in MNV and GA, whereby GA was additionally separated in paracentral, central, and central with preserved RPE-island (see Supplementary Table S1).

Acquisition of Genetic Data

Whole blood ascertained at baseline visit and stored at −20°C was used for isolation of genomic DNA by experienced staff, as described previously.[7,9] Precipitation methods (Gentra Puregene Blood Kit, Qiagen, Hilden, Germany; and modified therefrom) were applied and DNA was stored at −20°C in 10 mM Tris/1 mM EDTA buffer, pH 8.0.

Genotyping was conducted by the Genome Analysis Center, Helmholtz Zentrum Munich, Germany, using the Infinium Global Screening Array-24, GSAMD, version 1.0 (Illumina Inc., San Diego, CA, USA) for AugUR1, and by Life & Brain GmbH, Bonn, Germany (InfiniumGlobal Screening Array-24, GSA version 3.0, 10/2020) for AugUR2. The genotyped variant calling was conducted for 700,078 variants using the "GSA-24v1-0_A1_ClusterFile.egt" and for 730,059 variants using the "GSAMD-24v3-0-EA_20034606_A1" cluster files, respectively. The DNA samples were genotyped on 50 batches. No batch effect was observed. Variants with a variant call rate < 95% and not in Hardy-Weinberg equilibrium (HWE; $P < 5 \times 10^{-8}$) were excluded. After removing duplicated samples, we identified an individual's genetic ancestry by merging their genotypes with those from the Human Genome Diversity Project (HGDP)[11] and computing the first two principal components (PCs). We categorized individuals as being of European, Asian, African, or other genetic ancestry based on how they clustered with the same genetic ethnicity in the HGDP data. We estimated relatedness between individuals using the KING software,[12] declaring individuals as "related" if they were second degree related to at least one other individual in the study, and as "unrelated" if they were not.

Imputation and Genetic Risk Score Calculation

Genotyping, imputation, and GRS-computation was done separately for AugUR1 and AugUR2 restricted to unrelated and European individuals (AugUR1: $n = 1099$ and AugUR2: $n = 1225$; Supplementary Fig. S1). We generated genetic PCs based on the genotyped variants and used the first 10 PCs to capture population substructure in the genetic association analyses.

Before imputation, we aligned genotyped variants to match the TOP strand annotation of the 1000 Genomes Phase III version 5 reference data and excluded variants with allele frequency differences from the study and reference data. Imputation was conducted using the two-step approach of imputation.[13] We estimated study haplotypes with ShapeIT[14] and imputed untyped variants with minimac3.[15] The 47,109,465 imputed autosomal variants were coded as the estimated number of copies of the specified allele (allelic dosage).

Previously, 52 variants were identified by a genomewide association study for late AMD using logistic regression adjusted for PCs and evaluated also when adding age and sex into the genetic association model.[8] Of these 52 variants, 50 variants were available for analysis in AugUR1 and AugUR2 (Supplementary Table S2). The GRS was generated by multiplying each variant's individual allele dosage by its respective weight (i.e. effect sizes from Fritsche et al.[8]), and then summing across all 50 variants. The GRS was then scaled by dividing through the average weight: let k denote the number of trait associated variants j in one individual i with individual allele dosage ($0 \leq \text{dosage}_{ij} \leq 2$), and let b_j denote the weight of SNP j, then the GRS_i for person i can be written as follows:

$$GRS_i = \frac{\sum_{j=1}^{k} dosage_{ij} b_j}{\frac{1}{k}\sum_{j=i}^{k} b_j}$$

By this, one GRS unit reflected one risk allele of average effect.

Data Management and Statistical Analyses

Askimed (http://www.askimed.com/) and SAS 9.4 software (SAS Institute Inc., Cary, NC, USA) were used for data management. The GRS was computed using R version 4.2.2. Descriptive statistics and association analyses were carried out using IBM SPSS Statistics for Windows, version 29.0 (IBM Corp., Armonk, NY, USA). Figures were generated using R extended by ggplot 2.[16]

The analyzed sample yielded all AugUR participants with available genetic information and available AMD status in at least one eye. We here analyzed the AugUR data cross-sectionally using the genetic information and AMD status from the AugUR1 and the AugUR2 baseline surveys. For the replication of the previously observed pattern when comparing CC versus 3CACCSS classification, we restricted this to AugUR2.

For descriptive statistics, mean ± standard deviation was provided for continuous normally distributed variables and % (n) for dichotomous variables, unless stated otherwise.

In order to derive the association of the GRS with AMD of various stages or AMD features, we performed logistic regression in this cross-sectional data with the GRS as independent variable and AMD stages or AMD features as the dependent variable (primary analysis). As covariates, we included the first 10 PCs to account for genetic population substructure and a study variable for the AugUR survey membership to account for the design (AugUR2 versus AugUR1); the inclusion of these covariates had

little, if any, impact and were only added for methodological reasons. In the primary model, we derived the GRS-association with AMD adjusted for age and sex, which is typical for genetic association analyses, "AMD ~ age, sex, GRS, 10 PCs, and survey membership" ("full model", *model 1*).

We did not include further AMD risk factors like smoking or obesity into this model, because we were not interested in the GRS association on AMD independent of smoking or obesity, but in the general prediction of the GRS for AMD. We also applied a model without adjusting for age and sex, "AMD ~ GRS, 10 PCs, and survey membership" (*model 2*), and compared the GRS effect estimates from *model 2* with those from *model 1*.

In the next step, we wanted to quantify the discriminatory ability of the GRS together with age and sex, of the GRS alone, and we also wanted to compare this with age and sex alone. We thus calculated the area under the receiver operating characteristics curve (AUC) using predicted probabilities from the logistic regression analyses for the respective AMD stage or AMD-related feature using (i) *model 1*, (ii) *model 2*, and (iii) a third model only including age and sex ("AMD ~ age, sex, and survey membership", *model 3*).

Results

Participant Characteristics and Distribution of AMD Stages

Among the 2449 AugUR participants (AugUR1: $n = 1133$ and AugUR2: $n = 1316$; and no overlap between the 2 datasets), information on AMD stages for at least one eye, as well as genetic data were available for 2188 persons (AugUR1: $n = 1010$ and AugUR2: $n = 1178$). In this analyzed sample, age ranged from 70 to 95 years (mean age = 78.3 ± 5.0), 52% were women; and 167 individuals (7.6%) had late AMD (Table 1). As reported previously,[3] the frequencies of "no AMD" and "early AMD" stages differed substantially between classification systems (CC = 48.7% and 43.7% versus 3CACSS = 72.1% and 20.3%, respectively; see Table 1).

AugUR2 Data Replicates the Previously Observed Pattern When Comparing 3CACSS- and CC-Based Classification

When adding the 1178 AugUR2 participants to the comparison of the individuals' classification via CC versus 3CACSS, we confirmed the pattern previously observed in 1010 AugUR1 participants (Table 2; showing the two datasets separately in Supplementary Table S3). This supported the notion that CC "age-related changes" and CC "early AMD" were distinct from 3CACSS "early AMD" stages.

Of note, there had been few ($n = 2$) individuals in AugUR1 with AMD-related pigmentary abnormalities but without any drusen, which is "intermediate AMD" by CC and "no AMD" by 3CACSS. We re-evaluated the color fundus images of these two individuals for misclassification, but confirmed the initial grading. We considered these as "age-related changes" in a 7-step system.

This 7-step system enabled us (i) to differentiate the 446 individuals with CC-based "intermediate AMD" by 3CACSS "mild", "moderate", or "severe early" AMD stages, and (ii) to differentiate the 1575 individuals with 3CACSS-based "no AMD" into individuals without any AMD (no CC-based nor 3CACSS-based AMD) nor age-related changes, or individuals

Table 1. Participant Characteristics of Analyzed Study Sample

	All ($n = 2188$)	Women ($n = 1140$)	Men ($n = 1048$)
Age, years, mean ± SD (min. – max.)	78.3 ± 5.0 (70–95)	78.4 ± 5.0	78.1 ± 5.1
GRS*, mean ± SD (min. – max.)	44.5 ± 3.8 (33.3 –56.0)	44.5 ± 3.8	44.5 ± 3.8
Clinical Classification (CC)			
No AMD stages overall, % (*n*)	48.7 (1065)	43.2 (493)	54.6 (572)
No AMD, no age-related changes, % (*n*)	28.8 (631)	26.4 (301)	31.5 (330)
No AMD, age-related changes, % (*n*)	19.8 (434)	16.8 (192)	23.1 (242)
Early AMD stages overall, % (*n*)	43.7 (956)	48.9 (557)	38.1 (399)
Early AMD, % (*n*)	23.3 (510)	27.0 (308)	19.3 (202)
Intermediate AMD, % (*n*)	20.4 (446)	21.8 (249)	18.8 (197)
Three Continent AMD Consortium Severity Scale (3CACSS)			
No AMD, % (*n*)	72.1 (1577)	70.4 (803)	73.9 (774)
Early AMD stages overall, % (*n*)	20.3 (444)	21.7 (247)	18.8 (197)
Mild early AMD, % (*n*)	9.3 (203)	9.3 (106)	9.3 (97)
Moderate early AMD, % (*n*)	5.4 (119)	6.8 (78)	3.9 (41)
Severe early AMD, % (*n*)	5.6 (122)	5.5 (63)	5.6 (59)
Late AMD†			
Late AMD overall, % (*n*)	7.6 (167)	4.1 (90)	3.5 (77)
MNV, % (*n*)	4.3 (95)	4.5 (51)	4.2 (44)
GA, % (*n*)	1.6 (35)	1.8 (20)	1.4 (15)
MNV and GA, % (*n*)	1.7 (37)	1.7 (19)	1.7 (18)

The analyzed sample consisted of 2188 participants with gradable color fundus images for at least one eye and available genetic data for AugUR1 and 2 individuals. Shown are participants characteristics overall and by sex, including the proportion (and number) of individuals that have no AMD or early AMD stages based on the Three Continent AMD Consortium Severity Scale (3CACSS)[2] or the Clinical Classification (CC)[4] or any late AMD stages.

* Weighted GRS based on 50 variants known for late AMD,[8] with one unit being the average effect of included variants.

† Late AMD definition is similar between the 3CACSS[2] and the CC.[4]

SD = standard deviation; GRS = genetic risk score; AMD = age-related macular degeneration; MNV = macular neovascularization; GA = geographic atrophy.

Table 2. A Distinct Pattern When Comparing the Two Established AMD Classification Systems Using AugUR Study Participant Data

		Three Continent AMD Consortium Severity Scale					
		No AMD	Mild Early AMD	Moderate Early AMD	Severe Early AMD	Late AMD	Total, *n* (%)
Clinical Classification	**No AMD, no age-related changes**	631	0	0	0	0	631 (28.8)
	No AMD, age-related changes	434	0	0	0	0	434 (19.8)
	Early AMD	510	0	0	0	0	510 (23.3)
	Intermediate AMD	2	203	119	122	0	446 (20.4)
	Late AMD	0	0	0	0	167	167 (7.6)
	Total, *n* (%)	1577 (72.1)	203 (9.3)	119 (5.4)	122 (5.6)	167 (7.6)	2188 (100)

We analyzed all AugUR participants with gradable color fundus images for at least one eye and available genetic data for AugUR1 and 2 individuals ($n = 2188$). Shown are the numbers of participants by AMD status for the Three Continent AMD Consortium Severity Scale[2] and the Clinical Classification.[4] The percentages in the "total" rows and columns highlight the distribution of AMD severity groups. The same pattern as observed previously for AugUR1 participants (as published in Brandl et al., Sci Rep., 2018[3]) was observed when adding the second independent dataset of AugUR2 (see Supplementary Table S3).

AMD = age-related macular degeneration.

with age-related changes, or individuals with CC-based "early AMD".

Distribution of the GRS in the Population-Based Sample Separately For Late AMD Cases and CC-Based Controls

Of the 52 variants previously identified for late AMD from a large case-control sample,[8] 50 variants were available in the population-based AugUR study (2 variants were not genotyped nor imputed). We derived the 50-variant GRS for each individual (AugUR1 and 2; $n = 2188$) and compared the GRS distribution between late AMD individuals (embedded cases) with the AMD-free participants (no AMD, no age-related changes as defined by CC). We observed a clear shift in the distributions toward an increased number of risk alleles in individuals with late AMD and a significant association compared to AMD-free individuals (odds ratio [OR] = 1.41, P value < 0.0001; Supplementary Fig. S2). The mean values of the GRS increased by AMD stages for both classification systems (Supplementary Table S4).

GRS Association Supports the 7-Step AMD Classification Scale Combining CC and 3CACSS

We computed the GRS association for case groups of increased AMD severity (steps ii to vi of the 7-step scale) and control groups with increased restrictions (only step i or adding step ii or even step iii in AugUR1 and 2, $n = 2188$). First, we derived the GRS association independent of age and sex using logistic regression with age, sex, GRS, 10 PCs, and survey membership as covariates (*model 1*) and each of the different case and control group definitions as outcome (cases: age-related changes, $n = 434$; CC early, $n = 510$; 3CACSS mild, $n = 203$; 3CACSS moderate, $n = 119$; 3CACSS severe, $n = 122$; controls: CC "no AMD and no age-related changes", $n = 631$; CC "no AMD but age-related changes", n = 1065; 3CACSS "no AMD", $n = 1577$).

When looking at the ORs for the GRS on 3CACSS AMD stages as case groups and 3CACSS "no AMD" as the control group, we observed a stepwise increase in the ORs by using higher stages of 3CACSS "early AMD" (OR = 1.08, 95% confidence interval [CI] = 1.04, 1.13, OR = 1.29, 95% CI = 1.22, 1.37, or OR = 1.38, 95% CI = 1.29, 1.47 for mild, moderate, or severe early AMD, respectively; Fig. 1A; Supplementary Table S5A). Accordingly, the GRS association with CC "intermediate AMD" as the case group captured an average OR across the three above noted ORs (OR = 1.20, 95% CI = 1.16, 1.24; see Fig. 1A).

When computing the GRS association with CC "early AMD" or even with CC "age-related changes" as the case group (using CC "no AMD without age-related changes" as controls), ORs for CC "early AMD" or for 3CACSS "mild early AMD" were similar (OR = 1.14, 95% CI = 1.10, 1.18 or OR = 1.14, 95% CI = 1.09, 1.20, respectively). Interestingly, we found a small but still statistically significantly increased OR for age-related changes as the case group (OR = 1.05, 95% CI = 1.01, 1.09). Thus, the GRS association supports the stepwise increased association by severity of the case groups including age-related changes, except for a plateau between CC "early AMD" and 3CACSS "mild early AMD".

Remarkably, consistently for each case group, the GRS associations increased by increased restriction of controls: the ORs increased when restricting the control group from 3CACSS "no AMD" to CC "no AMD including age-related changes" and CC "no AMD without age-related changes" (see Fig. 1A; Supplementary Table S5A). Although we found little differences in the ORs upon exclusion or inclusion of the CC-age-related changes in the control group, the ORs were always the same or higher when excluding age-related changes. Together with the significant OR of 1.05 for age-related changes as case group, the GRS associations supported the restriction of the control group to CC-based "no AMD without age-related changes".

When comparing the ORs for the GRS associations on early AMD stages independent of age and sex (i.e. adjusted for age and sex, *model 1*) with the GRS associations unadjusted for age sex (i.e. *model 2*), we found little differences (see Supplementary Table S5A).

Diagnostic Accuracy of the GRS Increases by Increased Early AMD Stages

We further computed the AUC and evaluated the ability of the GRS to discriminate each AMD stage versus "no AMD" with increased severity of cases and increased restrictions of controls. For this, we used logistic regression in the same case and control groups described above. However, we focused now on GRS association without adjustment for age and sex (i.e. *model 2*), because we were primarily

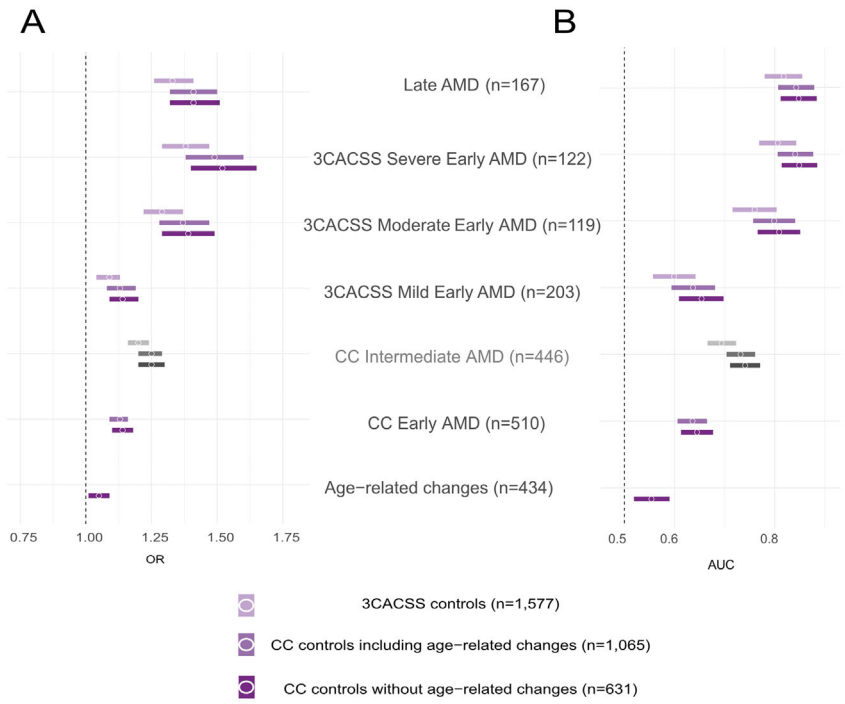

Figure 1. Association and diagnostic accuracy of the GRS on various AMD stages based on two established systems. We evaluated the 50-variant GRS for association with each AMD stage separately as outcome in all AugUR individuals with at least one eye gradable and genetic data available (n = 2188). Cases were defined as individuals in the respective AMD stage according to the 3CACSS[2] and the CC[4] systems (indicated vertically); color coding indicated the different control definitions (as applicable given the case group). We applied two logistic regression models "AMD ~ GRS, age, sex, 10PCs, and survey membership" (*model 1*) and "AMD ~ GRS, 10 PCs, and survey membership" (*model 2*). (**A**) Shown are odds ratios (ORs) and 95% confidence intervals (CIs) of the GRS association with stages of AMD using *model 1*. ORs were nearly the same when using *model 2* (see Supplementary Table S5A). (**B**) Diagnostic accuracy of the GRS on the for the various AMD stages is shown as area under the curve (AUC) and 95% CI based on predicted probabilities derived from *model 2* (see the details in Supplementary Table S5B). Intermediate AMD is displayed in grey due to the overlap with 3CACSS early AMD stages.

Table 3. AMD-Related Features in the AugUR Study Participants Overall and by Sex

Early AMD Features	All (n = 1966)	Women (n = 1037)	Men (n = 929)
Drusen			
Peripheral drusen only, % (n)	3.7 (73)	5.2 (54)	2.0 (19)
Small drusen only, % (n)	13.3 (261)	10.1 (105)	16.8 (156)
Intermediate drusen excluding peripheral drusen, % (n)	11.8 (232)	10.5 (109)	13.2 (123)
Intermediate drusen including peripheral drusen	18.7 (367)	20.8 (216)	16.3 (151)
Large drusen excluding peripheral drusen, % (n)	3.9 (76)	3.5 (36)	4.3 (40)
Large drusen including peripheral drusen, % (n)	8.0 (158)	9.5 (98)	6.5 (60)
Large drusen < O2 circle*, % (n)	4.6 (90)	5.0 (52)	4.1 (38)
Large drusen ≥ O2 circle*, % (n)	3.5 (68)	4.4 (46)	2.4 (22)
Subretinal drusenoid deposits only, % (n)	0.9 (19)	1.2 (12)	0.6 (7)
Pigment abnormalities			
Non-AMD related pigment abnormalities only, % (n)	4.6 (91)	3.2 (33)	6.2 (58)
Non-AMD related pigment abnormalities including small drusen, % (n)	7.0 (137)	4.6 (48)	9.6 (89)
AMD related pigment abnormalities including small and intermediate drusen, % (n)	1.9 (37)	1.5 (16)	2.3 (21)
AMD related pigment abnormalities including large drusen, excluding reticular drusen, % (n)	6.9 (135)	5.4 (56)	8.5 (79)
AMD related pigment abnormalities including large drusen, excluding peripheral drusen only, % (n)	11.6 (229)	11.5 (119)	11.8 (110)
Late AMD features			
MNV, % (n)	4.0 (79)	3.9 (40)	4.2 (39)
GA, % (n)	1.5 (29)	1.7 (18)	1.2(11)
Paracentral GA, % (n)	0.5 (10)	0.5 (5)	0.5 (5)
Central GA with preserved island, % (n)	0.15 (3)	0.3 (3)	0 (0)
Central GA, % (n)	0.7 (13)	0.8 (8)	0.5 (5)

We evaluated each AMD-related feature detectable on color fundus images which are part of the Clinical Classification or Three Continent AMD Consortium Severity Scale, plus peripheral drusen, subretinal drusenoid deposits, and non-AMD related pigment abnormalities. For this, we restricted the sample to the number of 1966 of AugUR participants with both eyes gradable. Shown is the proportion (number) of individuals with each feature overall and by sex.

* O2 circle is defined as a circle with diameter of 650 µm; drusen area equivalent to O2 circle accounts to 331.820 µm^2.

AMD = age-related macular degeneration; MNV = macular neovascularization; GA = geographic atrophy.

Investigative Ophthalmology & Visual Science

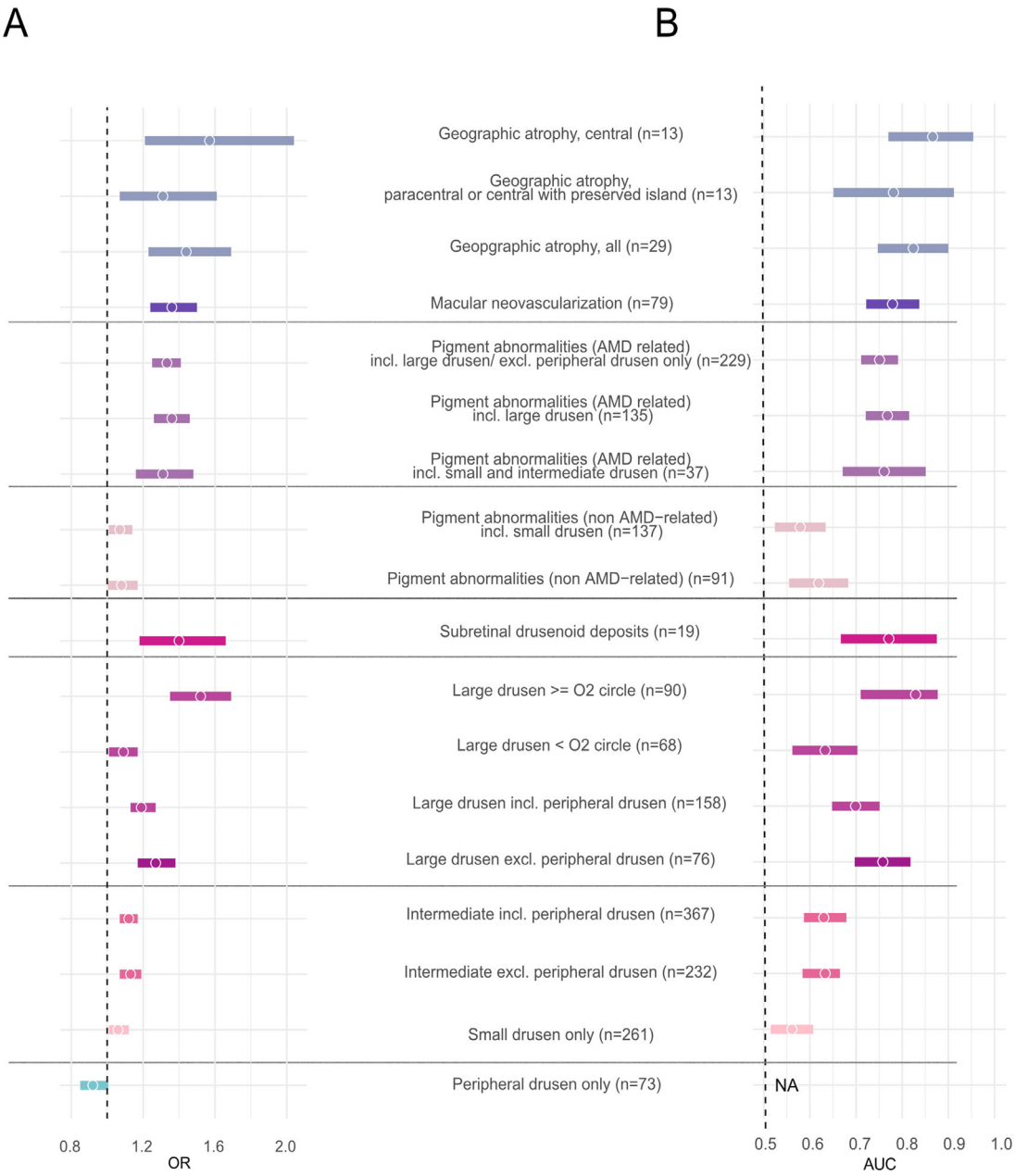

Figure 2. Association and diagnostic accuracy of GRS on AMD-related features. We evaluated the 50-variant GRS association with each AMD features separately as outcome restricting the analysis to AugUR individuals where both eyes were gradable (overall $n = 1966$). Cases were defined as individuals with the respective feature in at least one eye and controls were fully free of any feature (number of controls = 354): We applied two logistic regression models "AMD-feature ~ GRS, age, sex, 10 PCs, and survey membership" (*model 1*) and "AMD-feature ~ GRS, 10 PCs, and survey membership" (*model 2*). (**A**) Shown are odds ratios (ORs) and 95% confidence intervals (CIs) of the GRS association with AMD-related features using *model 1*. ORs were nearly the same when using *model 2* (see Supplementary Table S6A). (**B**) Diagnostic accuracy of the GRS for the various AMD stages is shown as area under the curve (AUC) and 95% CI based on predicted probabilities derived from *model 2* (see the details in Supplementary Table S6B). Color code was used to differentiate into groups of related features. AUC for peripheral drusen indicates not applicable (*NA*), because the GRS association estimate for peripheral drusen as outcome was negative and the AUC <0.5.

interested in the discriminatory ability of the GRS alone. We found a stepwise increase of the AUC when using 3CACSS mild, moderate, severe early, or late AMD as case groups and 3CACSS "no AMD" as controls (AUC = 58.3%, 95% CI = 54.0, 62.6, AUC = 73.2%, 95% CI = 68.7, 77.6, AUC = 77.9%, 95% CI = 74.0, 81.9, and AUC = 75.2%, 95% CI = 71.4, 79.1, respectively; see Fig. 1B; Supplementary Table S5B).

These AUCs increased, for each of the case groups, the more we restricted the control group (Fig. 1B; see Supplementary Table S5B). This resulted in a stepwise increase of the AUCs for each case group in the 7-step scale (CC "age-related chances", CC "early", 3CACSS "mild early", 3CACSS "moderate early", and 3CACSS "severe early", late AMD) using CC "no AMD without age-related changes" as controls: AUC = 55.1%, 95% CI = 51.6, 58.6, AUC = 62.3%, 95% CI =

59.1, 65.6, AUC = 63.8%, 95% CI = 59.3, 68.3, AUC = 78.1%, 95% CI = 73.6, 82.5, AUC = 82.2%, 95% CI = 78.4, 86.0, and AUC = 79.2%, 95% CI = 75.4, 83.0, respectively.

We observed a similar stepwise increasing pattern of the AUCs when including age and sex into the model (*model 1*; see Supplementary Table S5B), which we visualized also together with the AUC for age and sex alone (*model 3*; Supplementary Fig. S3).

Association of GRS With Separate AMD-Related Features

The AMD classification systems are based on combinations of AMD features detected on color fundus images. In order to understand the above noted findings on a more granular level, we evaluated each of these AMD features and combinations restricting the analyzed sample to individuals with both eyes gradable for AMD (n = 1966). Additionally, we investigated peripheral drusen, SDDs, and non-AMD-related pigmentary abnormalities, which are not part of the two established classification systems. The most frequent AMD-related features were small drusen, intermediate drusen, or large drusen (approximately 70% of individuals have at least one of these; Table 3).

We derived the GRS association with each of these features as outcome by defining individuals with the respective feature as cases and individuals without any detectable feature as controls (number of controls = 354). For this, we used logistic regression with GRS, age, sex, 10 PCs, and study membership as covariates (*model 1*). The GRS showed no significant association with peripheral drusen (OR = 0.92, 95% CI = 0.85, 1.006; Fig. 2A, Supplementary Table S6A), but increased association by increased drusen size (OR = 1.06, 95% CI = 1.01, 1.12, OR = 1.13, 95% CI = 1.07, 1.19, or OR = 1.27, 95% CI = 1.17, 1.38 for small, intermediate, or large drusen, respectively). Significant associations were also observed for drusen area with a higher OR for large drusen area compared to smaller drusen area (OR = 1.51, 95% CI = 1.35, 1.69 or OR = 1.09, 95% CI = 1.01, 1.17 for drusen $\geq$ O2 circle or $<$ O2 circle, respectively). SDDs also showed high GRS association (OR = 1.40, 95% CI = 1.18, 1.66). The strongest GRS association was shown with central GA (OR = 1.57, 95% CI = 1.21, 2.04). The ORs for the GRS with AMD features were similar without adjustment for age and sex (*model 2*; see Supplementary Table S6A).

We computed the diagnostic ability of the GRS for each feature as AUC based on the logistic regression model without age and sex adjustment (*model 2*). Differentiating the drusen area of large drusen yielded a higher diagnostic accuracy of the GRS for drusen $\geq$ O2 circle (AUC = 82.9%, 95% CI = 71.0, 87.7%) compared to drusen $<$ O2 (AUC = 63.3%, 95% CI = 56.2, 70.3; see Fig. 2B, Supplementary Table S6B). SDDs again also showed high values for AUC (77.1%, 95% CI = 66.7, 87.5). We found the highest diagnostic accuracy for central GA (AUC = 86.6%, 95% CI = 77.0, 95.4%).

For each AMD feature, the AUC of the GRS alone was higher than the AUC of age and sex (except for "peripheral drusen only"), but lower than the AUC of the GRS together with age and sex (Supplementary Fig. S4).

Discussion

In our analysis, we used the genetics of late AMD to support the utility of a 7-step AMD severity scale, which merged the CC and 3CACSS classification systems. We demonstrate a refinement on both ends of the scale when these two systems are combined: a refinement of "early AMD" and a refinement of the control group.

Our cross-sectional data of approximately 1200 individuals aged 70 to 95 years from our population-based AugUR2 study and the comparison of their CC versus 3CACSS-based AMD classification confirmed the distinct pattern previously observed in the AugUR1 data of 1010 non-overlapping individuals.[3] Specifically, the 3CACSS-based "mild/moderate/severe early AMD" differentiated the CC-based "intermediate AMD", due to 3CACSS additionally measuring and differentiating drusen area (see Supplementary Table S1). In addition, the 3CACSS-based "no AMD" is differentiated by the CC "age-related changes" and CC "early AMD", because intermediate drusen (>63 µm and $\leq$125 µm diameter) without any AMD-related pigmentary abnormalities are recorded as "early AMD" in CC but as "no AMD" in 3CACSS (see Supplementary Table S1). Our results thus support our idea of a 7-step scale that we propose to term: (i) "no AMD", (ii) "age-related changes", (iii) "very early AMD", (iv) "mild early AMD", (v) "moderate early AMD", (vi) "severe early AMD", and (vii) "late AMD" (see Supplementary Table S1). This system has the immediate benefit of combining the information of two established systems and by providing an unambiguous naming scheme for early stages of AMD pathology. The value of this merged system was emphasized by GRS association analyses which demonstrated a stepwise increased discriminatory ability when using the early AMD steps ii to vi as case groups and when excluding the steps ii and iii from the control group. This supports the utility of the 7-step scale.

The utility of this newly proposed 7-step AMD severity scale is also defined by the convertibility of already applied AMD classification approaches into this new system – preferably without regrading the entire study sample. Epidemiological studies on AMD which use the CC would have to regrade only individuals/eyes with "intermediate AMD" to subdivide them into "mild/moderate/early AMD": they would have to measure drusen area $<$ or $\geq$ O2 circle.[17] Studies applying the Rotterdam Classification would also have to measure drusen area $<$ or $\geq$ O2 circle in participants/eyes with "soft indistinct drusen $\geq$125 µm".[18] Studies applying the Age-related Eye Disease Study (AREDS) simplified severity scale would have to assess small and intermediate drusen (if not already done in the grading process anyway) and would also have to measure drusen area $<$ or $\geq$ O2 circle in participants with "large drusen".[5] The AREDS 9-step severity scale and the Wisconsin Age-related Maculopathy Grading System, however, assess all required features in more details as needed for the 7-step scale and can be converted by an algorithm without regrading.[19,20]

It is compelling to use genetic risk variants to quality control a clinical AMD classification system given the strong genetic risk of late AMD.[8] Our previous work has used genetic risk to quality control machine learning based automated AMD classification.[21] We here extend this idea to evaluate the utility of our proposed 7-step scale. This has the limitation that AMD pathways beyond age, sex, and the genetics captured by the known 50 risk variants are disregarded here. However, the genetic risk variants have been identified by a large genomewide association study with >17,000 late AMD cases and respective controls.[8] Furthermore, the predictive ability of the 50-variant GRS, age, and sex reached 84% for late AMD versus the most restricted

control group, which documented the substantial discriminatory ability.

The "classic" way to establish and quality control a classification system for early disease stages is to document the performance to predict late AMD in a stepwise fashion using individuals with early AMD and longitudinal data on progression. However, such longitudinal data are scarce and have not yet sufficiently answered the search for the "best" classification system.[6] Using a GRS for late AMD and investigating its association and diagnostic accuracy for early AMD stages is a compelling alternative. It allows for using the larger abundance of cross-sectional data – because genetics do not change over time or by disease status. To our knowledge, our study represents the first approach to utilize the large genetic heritability of AMD to help evaluate AMD classification systems. We do acknowledge the limitation that genetics is not the perfect comparator and that validation via longitudinal analyses is still warranted.

In clinical routine, AMD features are typically diagnosed via funduscopy and multimodal imaging – and they are interpreted individually without computing them into classification systems.[22,23] Therefore, we also investigated the single AMD-related features separately and found increased GRS association and diagnostic accuracy by increased drusen size and area. Large drusen with a large drusen area, SDDs, and central GA showed the highest GRS association and diagnostic accuracy. Particularly for SDDs, this is also in line with literature substantiating the risk of SDDs for AMD progression.[24] Interestingly, it seems that peripheral drusen do not share the genetics of late AMD.

One has to keep in mind that, in our analyses, AMD-phenotyping and classification are based on color fundus imaging, as this is still the current gold standard of AMD definition in epidemiological studies.[6] However, particularly SDDs are known to be better visualized with other retinal imaging modalities.[24] Of note, there is no established AMD classification system based on multimodal imaging to date, including, for example, optical coherence tomography. The European Eye Epidemiology Consortium has started a first approach[25]; further studies as to how multimodal imaging can improve AMD classification are warranted – including predictive ability analyses with progression data. Joint international efforts of epidemiological studies to find a consensus on a robust, unified system of AMD grading are crucially important to distinguish particularly the early stages of AMD and differentiate them from normal aging. The AugUR study, with its follow-up recruitment being completed in 2025, will be able to contribute to these efforts in the near future. Due to the high age of our study participants, we will be able to observe high numbers of different early AMD stages and features.

In summary, we demonstrate that combining the CC and 3CACSS to a 7-step AMD classification system is helpful in future large scale epidemiological studies on clarifying AMD causation. We used genetics from a population-based study of elderly individuals to evaluate its performance. The association as well as the diagnostic accuracy of the GRS stepwise increased by severity of AMD disease stage and was more significant when using a more restrictive definition of "no AMD". The combined 7-step scale provides an immediate improvement over using either CC or 3CACSS by using the full information and by sharpening the control group. It is also an immediate solution to the currently confusing terminology regarding "early AMD". Studies with existing AMD classification only need to apply a converting algorithm or regrade specific features in a subgroup, for which we provide a practical guide. We thus suggest the proposed 7-step scale as a consensus for AMD classification and recommend implementing it in epidemiological studies.

Acknowledgments

The authors gratefully thank the excellent supporting assistance of Lydia Mayerhofer, Josef Simon, and Sylvia Pfreintner. Moreover, we thank all study participants for contributing to the AugUR study.

The AugUR study and analyses are supported by grants from the German Federal Ministry of Education and Research (BMBF 01ER1206, BMBF 01ER1507 to I.M.H.), by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation; HE 3690/7-1 and HE 3690/5-1 to I.M.H., BR 6028/2-1 to CB), by the National Institutes of Health (NIH R01 EY RES 511967 and 516564 to I.M.H.), and institutional budget (University of Regensburg). The sponsors or funding organizations had no role in the design or conduct of this research.

Author Contributions: J.M.H.: manuscript writing, calculation of genetic risk score, and statistical analyses; M.E.Z.: data management, statistical analyses; M.G.: imputation of genetic data; F.G.: statistical analyses; BHFW: AugUR study co-principal investigator, manuscript writing; H.H.: manuscript writing, ophthalmological expertise; K.J.S.: AugUR study coordinator, project supervision, manuscript writing, statistical analyses, interpreting results; I.M.H.: AugUR study principal investigator, project initiation and supervision, manuscript writing, interpreting results; C.B.: project initiation and supervision, eye sub-study coordinator for AugUR, manuscript writing, interpreting results, grading of color fundus images. All authors contributed to the reviewing and editing of the manuscript.

Disclosure: **J.M. Herold**, None; **M.E. Zimmermann**, None; **M. Gorski**, None; **F. Günther**, None; **B.H.F. Weber**, None; **H. Helbig**, Alcon (R), Allergan (R), Apellis (R), Bayer (R), Novartis (R), and Theapharm (R); **K.J. Stark**, None; **I.M. Heid**, Roche Diagnostics (R) for a project related to AMD, but unrelated to this work presented here; **C. Brandl**, None

References

1. Lim LS, Mitchell P, Seddon JM, Holz FG, Wong TY. Age-related macular degeneration. *Lancet*. 2012;379(9827):1728–1738.
2. Klein R, Meuer SM, Myers CE, et al. Harmonizing the classification of age-related macular degeneration in the three-continent AMD consortium. *Ophthalmic Epidemiol*. 2014;21(1):14–23.
3. Brandl C, Zimmermann ME, Gunther F, et al. On the impact of different approaches to classify age-related macular degeneration: results from the German AugUR study. *Sci Rep*. 2018;8(1):1–10.
4. Ferris FL, 3rd, Wilkinson CP, Bird A, et al. Clinical classification of age-related macular degeneration. *Ophthalmology*. 2013;120(4):844–851.
5. Ferris FL, Davis MD, Clemons TE, et al. A simplified severity scale for age-related macular degeneration: AREDS Report No. 18. *Arch Ophthalmol*. 2005;123(11):1570–1574.
6. Thee EF, Meester-Smoor MA, Luttikhuizen DT, et al. Performance of classification systems for age-related macular degeneration in the Rotterdam Study. *Transl Vis Sci Technol*. Published online 2020.
7. Brandl C, Günther F, Zimmermann ME, et al. Incidence, progression and risk factors of age-related macular degen-

eration in 35–95-year-old individuals from three jointly designed German cohort studies. *BMJ Open Ophthalmol*. 2022;7(1):e000912.
8. Fritsche LG, Igl W, Bailey JNC, et al. A large genome-wide association study of age-related macular degeneration highlights contributions of rare and common variants. *Nat Genet*. 2016;48(2):134–143.
9. 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;15:130.
10. Brandl C, Zimmermann ME, Herold JM, Helbig H, Stark KJ, Heid IM. Photostress recovery time as a potential predictive biomarker for age-related macular degeneration. *Transl Vis Sci Technol*. 2023;12(2):15.
11. Cavalli-Sforza LL. The Human Genome Diversity Project: past, present and future. *Nat Rev Genet* 2005;6(4):333–340.
12. Manichaikul A, Mychaleckyj JC, Rich SS, Daly K, Sale M, Chen WM. Robust relationship inference in genome-wide association studies. *Bioinformatics*. 2010;26(22):2867–2873.
13. Howie B, Fuchsberger C, Stephens M, Marchini J, Abecasis GR. Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. *Nat Genet*. 2012;44(8):955–959.
14. Delaneau O, Zagury JF, Marchini J. Improved whole-chromosome phasing for disease and population genetic studies. *Nat Methods*. 2013;10(1):5–6.
15. Das S, Forer L, Schönherr S, et al. Next-generation genotype imputation service and methods. *Nat Genet*. 2016;48(10):1284–1287.
16. Villanueva RAM, Chen ZJ. *ggplot2: Elegant Graphics for Data Analysis* (2nd ed.). New York, NY; Springer: 2019;17(3):160–167.
17. Ferris FL, 3rd, Wilkinson CP, Bird A, et al. Clinical classification of age-related macular degeneration. *Ophthalmology*. 2013;120(4):844–851.
18. Klaver CC, Assink JJ, van Leeuwen R, et al. Incidence and progression rates of age-related maculopathy: the Rotterdam Study. *Invest Ophthalmol Vis Sci*. 2001;42(10):2237–2241.
19. Davis MD, Gangnon RE, Lee LY, et al. The Age-Related Eye Disease Study severity scale for age-related macular degeneration: AREDS Report No. 17. *Arch Ophthalmol*. 2005;123(11):1484–1498.
20. Klein R, Davis MD, Magli YL, Segal P, Klein BE, Hubbard L. The Wisconsin age-related maculopathy grading system. *Ophthalmology*. 1991;98(7):1128–1134.
21. Guenther F, Brandl C, Winkler TW, et al. Chances and challenges of machine learning-based disease classification in genetic association studies illustrated on age-related macular degeneration. *Genet Epidemiol*. 2020;44(7):759–777.
22. Spaide RF, Jaffe GJ, Sarraf D, et al. Consensus nomenclature for reporting neovascular age-related macular degeneration data: consensus on neovascular age-related macular degeneration nomenclature study group. *Ophthalmology*. 2020;127(5):616–636.
23. Spaide RF, Ooto S, Curcio CA. Subretinal drusenoid deposits AKA pseudodrusen. *Surv Ophthalmol*. 2018;63(6):782–815.
24. Wu Z, Fletcher EL, Kumar H, Greferath U, Guymer RH. Reticular pseudodrusen: a critical phenotype in age-related macular degeneration. *Prog Retin Eye Res*. 2022;88:101017.
25. Gattoussi S, Buitendijk GHS, Peto T, et al. The European Eye Epidemiology spectral-domain optical coherence tomography classification of macular diseases for epidemiological studies. *Acta Ophthalmol*. 2019;97(4):364–371.

Investigative Ophthalmology & Visual Science