Published in final edited form as:

Ophthalmology. 2013 September ; 120(9): 1880–1892. doi:10.1016/j.ophtha.2013.02.007.

Inclusion of Genotype with Fundus Phenotype Improves Accuracy of Predicting Choroidal Neovascularization and Geographic Atrophy

Lorah T. Perlee, PhD¹, Aruna T. Bansal, PhD², Karen Gehrs, MD³, Jeffrey S. Heier, MD⁴, Karl Csaky, MD⁵, Rando Allikmets, PhD⁶, Paul Oeth, MS⁷, Toni Paladino, PhD⁷, Daniel H. Farkas, PhD⁸, P. Lyle Rawlings, BS⁸, and Gregory S. Hageman, PhD⁹

¹Sequenom, Inc., San Diego, CA 92121, USA

²Acclarogen Limited, Cambridge CB4 0WS, UK

³Department of Ophthalmology and Visual Sciences, University of Iowa, Iowa City, IA 52242, USA

⁴Ophthalmic Consultants of Boston, Boston, MA 02114, USA

⁵Texas Retina Associated, Dallas, TX 75231, USA

⁶Departments of Ophthalmology and Pathology and Cell Biology, Columbia University, New York, NY 10032, USA

⁷Sequenom Center for Molecular Medicine, LLC (SCMM), San Diego CA, 92121, USA

⁸Sequenom Center for Molecular Medicine, LLC (SCMM), Grand Rapids, MI 49503, USA

⁹Department of Ophthalmology and Visual Sciences, John A. Moran Eye Center, University of Utah, Salt Lake City, UT 84132, USA

Abstract

Purpose—The accuracy of predicting conversion from early stage age-related macular degeneration (AMD) to the advanced stages of choroidal neovascularization (CNV) and/or

© 2013 American Academy of Ophthalmology, Inc. Published by Elsevier Inc. All rights reserved.

Address for Reprints: Lorah T. Perlee, PhD, Sequenom, Inc., 3595 John Hopkins Drive, San Diego, California U.S.A. 92121.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Meeting Presentation: American Academy of Ophthalmology Annual Meeting, November, 2012

Conflicts of Interest:

LTP, PO, TP, DHF, PLR: Sequenom, SCMM Employees

LTP, PO, TP, DHF, PLR: Sequenom Stock Options

ATB: Consulting Fee, Fees for Statistical Analysis, Payment for Writing

KG: Consulting Sequenom, Travel support, Fees, Board Membership, Consultancy other, Speaker Bureau, Payment lectures, Payment Presentations, Travel Expenses

JH: Consulting Sequenom, Consulting Other, Grants

KC: Consulting other

RA: Grants, Columbia University Employee, Grants other, Patents, Royalties

GSH: Consulting Sequenom, Board membership, Consultancy other, University Utah Employee, Speaker Bureau, Payment lectures, Patents, Royalties, Stock

Online Only: The following should appear online-only: Tables 4, 8, 9, 10, 13, 14, 15, 16, 18; Figures 1, 6, 9; Appendix 1

geographic atrophy (GA) was evaluated to determine if inclusion of clinically relevant genetic markers improved accuracy beyond prediction using phenotypic risk factors alone.

Design—Cohort study.

Participants—White, non-Hispanic subjects participating in the Age Related Eye Disease Study (AREDS) sponsored by the National Eye Institute, consented to provide a genetic specimen. Of 2,415 DNA specimens available, 940 were from disease-free subjects and 1,475 were from subjects with early or intermediate AMD.

Methods—DNA specimens from study subjects were genotyped for 14 single nucleotide polymorphisms (SNPs) in genes shown previously to associate with CNV: *ARMS2*, *CFH*, *C3*, *C2*, *FB*, *CFHR4*, *CFHR5* and *F13B*. Clinical demographics and established disease associations, including age, sex, smoking status, body mass index, AREDS treatment category, and educational level were evaluated. Four multivariate logistic models [Phenotype; Genotype; Phenotype + Genotype; Phenotype + Genotype + Demographic + Environmental factors] were tested using two endpoints (CNV, GA). Models were fitted using Cox proportional hazards regression to utilize time-to-disease onset data.

Main Outcome Measures—Brier score (measure of accuracy) was employed to identify the model with the lowest prediction error in the training set. The most accurate model was subjected to independent statistical validation and final model performance described using area under the receiver operator curve (AUC) or C-statistic.

Results—The CNV prediction models that combined genotype with phenotype with or without age and smoking revealed superior performance (C-statistic=0.96) compared to the phenotype model based on simplified severity scale and the presence of CNV in the non-study eye (C-statistic=0.89), p value < 0.01. For GA, the model that combined genotype with phenotype demonstrated the highest performance (AUC=0.94). Smoking status and *ARMS2* genotype had less of an impact on the prediction of GA compared to CNV.

Conclusions—Inclusion of genotype assessment improves CNV prediction beyond that achievable with phenotype alone and may improve patient management. Separate assessments should be used to predict progression to CNV and GA since genetic markers and smoking status do not equally predict both endpoints.

Age-related macular degeneration (AMD) is the leading cause of irreversible blindness, with an estimated 1.75 million United States patients affected with the late stage, vision-threatening form and another 7 million at high risk of developing advanced disease.¹ Reports indicate 64% of White patients diagnosed with advanced disease have choroidal neovascularization (CNV), an advanced form of the disease characterized by the presence of subretinal or sub-retinal pigment epithelial neovascularization, serous or hemorrhagic sensory retinal detachment, and/or sub-retinal pigment epithelial (RPE) hemorrhage.² Geographic atrophy (GA), the second most prevalent form of advanced disease (38%), manifests with areas of partial or complete depigmentation, atrophy, and loss of the RPE.² Although CNV impacts <10% of the AMD disease population, it accounts for >90% of the vision loss associated with AMD.³ Although no clinical intervention is currently available to reduce progression to GA, treatment regimens using anti-vascular endothelial growth factor (anti-VEGF) therapies have proven effective in reducing the near-term, CNV-associated vision loss in some patients. Importantly, timely diagnosis remains central to minimizing the irreversible vision-loss associated with treatment delay.⁴⁻⁶

Models developed to predict advanced stages of AMD are well documented.⁷⁻¹⁵ Genetic, phenotypic, demographic and/or environmental criteria associated with progressive disease have been used to identify patients at high risk of advanced AMD.⁷⁻¹⁵ Although earlier works have confirmed the contribution of genetics in risk prediction models, the magnitude

of the contribution was not deemed substantial. These earlier predictive models might have underestimated the contribution of genetics as they were not designed to resolve the haplotype assignments associated with the complex genetic structure of the complement factor H (CFH) and CFH-related protein gene cluster on chromosome 1.¹⁶ Hughes *et al.*¹⁶ first reported a predictive model for wet AMD based on CFH haplotypes and cautioned against the incomplete and inaccurate representation of this region that would result from models lacking the necessary SNPs to provide resolution of the five major CFH haplotypes. Hughes was also first to describe a protective haplotype containing a 85kb deletion in the CFHR1/3 region that conferred significant protection from disease compared to the increased susceptibility associated with the major risk haplotype reported by Hageman in 2005.^{17,18} Specifically, Hughes findings indicated “only haplotype analysis can assess the effect of combinations of all the coding variants in CFH that may contribute to AMD susceptibility.”¹⁶

Earlier studies have also been directed at the clinically meaningful endpoint of advanced disease that includes both GA and CNV. A recent study evaluating the genetic markers associated with progression to advanced AMD in Caucasians revealed a differential contribution of certain SNPs in AMD-related genes influencing the development of CNV compared to progression to GA.¹⁹ A similar observation was reported when comparing progression to CNV versus polypoidal choroidal vasculopathy (PCV) in an Asian cohort.²⁰ If multiple mechanisms are involved in the biology of the two major forms of advanced disease (CNV, GA), it will be important to develop independent models to predict the two distinct phenotypes until such time that additional discovery can predict whether a patient will progress to CNV and/or GA. Combining the two phenotypic endpoints into a single disease progression model may reduce the accuracy of prediction and or translate into differences in performance between single versus dual endpoint models.

Finally, there may be differences in the accuracy of prediction based on the fundus grading system applied in a given study. Several studies have utilized the AREDS grading system while others have relied on the use of the AREDS simplified severity scale to assess risk of progression. Baseline disease grading systems that capture the clinical indicators of progression from early-and intermediate-stage disease to advanced forms of disease have been reported. A simplified grading system (AREDS Simplified Severity Scale), based on the detection of one or more large drusen ($>125\ \mu\text{m}$) and the presence of any abnormal pigmentation, was developed to stratify patients at risk of progressing to advanced stage AMD.²¹ The risk factors are summed across both eyes, generating a 5-point scale (0–4) to estimate the 5-year risk of developing advanced AMD. The five-point scale is a strong predictor of advanced disease that is easy to use; thus it is often the only indicator, currently in broad use, employed to predict a patient’s risk of developing advanced disease. Prediction of a patient’s likelihood of progression to late stage disease is more often limited to an assessment of fundus phenotype due to a lack of consensus on the clinical value of including genetics into clinical care.²²

The purpose of this study was to determine if the inclusion of known, clinically relevant genetic markers further improves risk stratification of patients at high risk of developing CNV or GA compared to earlier models that predict advanced disease based on phenotypic, demographic and environmental risk factors alone. Progression to CNV and GA was surveyed using four different prediction models: a phenotype model that assessed simplified severity scale and advanced disease in the fellow eye; a genotype model that surveyed 14 disease-associated SNPs and advanced disease in the fellow eye; a combined model that included the assessment of phenotype with genotype; and a full model that combined the phenotype and genotype models with clinical and environmental variables. We isolated these variables to demonstrate the clinical significance of including genotyping for

accurately identifying patients at the highest risk of developing CNV or GA and demonstrated inclusion of genetic data supports the best possible patient management. These data substantiate the adoption of a genetic test that provides the managing physician with a percent likelihood of progression to advanced disease at practical time points and encourages more frequent surveillance of high risk patients to favor more timely diagnosis. By including genetic progression markers into the risk assessment, patient surveillance can be tailored to meet the needs of the individual.

Methods

Subjects

Data were derived from subjects participating in the Age-Related Eye Disease Study (AREDS). The AREDS trial was a multicenter, prospective, longitudinal study evaluating the clinical course of AMD and cataracts, as well as the effect of high dose vitamin/mineral supplementation on progression of these diseases.²³ Clinical, demographic, and environmental data for each participant were retrieved from the AREDS database of Genotype and Phenotype (dbGaP). The baseline disease assignment utilized in this study was based on the AREDS five-step (0–4) simplified severity scale²¹ with annual visit data graded according to the AREDS 12-point severity scale.²⁴ There were 603 subjects with a recorded CNV event and 379 subjects with a reported assignment of central GA in the dbGaP database during the follow-up period. DNA samples from 2,415 White, non-Hispanic, consented study subjects, ages 55–81, were obtained from the National Eye Institute-AREDS Genetic Repository located at the Coriell Institute for Medical Research (<http://ccr.coriell.org/Sections/Collections/AREDS/?SsId=68>, accessed February 4, 2013).²³ 940 DNAs were from disease-free subjects and 1,475 DNAs were obtained from subjects diagnosed with early or intermediate stages of AMD at baseline. The phenotypic data from the dbGaP tables were merged with the genetic data generated by conducting genotyping on the 2,415 AREDS DNA specimens with the 14 disease-associated SNPs.

This study applied the same definition of “progressors” used in the AREDS trial. The term, “progressors,” was defined as individuals with no, early or intermediate AMD at baseline who progressed to advanced AMD during follow-up and individuals with advanced AMD in one eye at baseline who progressed to advanced AMD in both eyes. The definition of a control was equivalent to the designation “non-progressor,” that was used to identify subjects with early or intermediate AMD that did not progress to CNV or GA during the follow-up period. Subjects spanning the entire range of the baseline simplified severity scale were analyzed, with an adjustment made for the presence of advanced disease in the non-study eye. During the clinical trial portion of the AREDS study, subjects were randomized into four treatment arms: antioxidants only, antioxidants with zinc, zinc only, and placebo. AREDS subjects fell into two consent categories, General Research Use (GRU) and Eye Disease Only (EDO).²³ The majority of subjects fell under the GRU category with just over 10% of subjects consented as EDO; for the latter, body mass index (BMI) data were not available.

Genetic Markers

Thirteen genetic markers (SNPs), generated by genotyping the 2,415 subject DNAs, were evaluated in the modeling and validation analyses (Table 1). Genotyping was also conducted on one additional SNP, rs11200638 in HTRA1, to evaluate its performance as a proxy for rs10490924 in ARMS2 to determine if clinical performance of the two markers was equivalent. Details about the performance of rs11200638 can be found in Appendix 1, available at <http://aaojournal.org>.

Training and Validation Sets

A training set was prepared using two-thirds of the subjects randomly sampled from the complete set of participants studied at baseline. A two-thirds split was based on power calculations demonstrating good power (>80%) for the detection of association to CNV for the major gene variants including rs641153 in CFB, that revealed a low frequency combined with a moderate effect. The balance of samples made up the validation set. Case-control matching was not conducted in order to maintain the prevalence rates in both phases and to ensure realistic estimates of model-performance. Table 2 shows the distribution of subjects across the training and validation sets.

Demographic Characteristics

Demographic characteristics were tabulated separately for progressor and non-progressor subjects for both CNV and GA. The subject category counts and mean values $\pm$ standard errors (SE) are presented in Table 3.

Quality Assessment by Subject

Each of the 2,415 subjects was genotyped for the 14 SNPs listed in Table 1 as previously described.⁹

Quality Assessment by Marker

The percentage of missing data and the minor allele frequency in non-CNV-progressors were calculated and tabulated for each marker. No markers were excluded on the basis of missing data, low minor allele frequency or departure from Hardy-Weinberg equilibrium. Full details on genotype counts, percent missing data, non-CNV minor allele frequency and non-CNV Hardy-Weinberg equilibrium can be found in Table 4 (available at <http://aaojournal.org>).

Model Development

Logistic regression models do not accommodate the time-to-onset of advanced disease, which varies among subjects, so the data were modeled as right-censored survival data using a Cox Proportional Hazards Model. This methodology used the time to onset (time-to-event) of advanced disease to establish the probability of advanced AMD over time. AMD status at each follow-up visit was defined using the AREDS 12-point AMD severity scale for each study eye. Advanced AMD was represented as Central GA (GA), CNV, or both GA and CNV in the same eye.²⁴

The individual variables influencing disease progression were assessed. The effect sizes of each of the phenotypic, genetic, demographic and environmental factors were estimated singly for each endpoint (CNV, GA). Next, four multivariate logistic models were evaluated, with the presence or absence of baseline advanced disease in a non-study eye included as the indicator variable in each of the four models.

Estimated hazard ratios (HR), natural log of standard error (ln (SE)), 95% confidence intervals (CI) and p-values were obtained for each variable in each of the four models tested. The performance of the four models was compared using a 10-fold cross-validation estimate of Brier Score or mean square error (MSE). For each endpoint, the model with the lowest estimated error (MSE) was identified and selected for validation. The Brier score was predefined in the analysis plan to evaluate the prediction error in the training data set observed in the four predictive models to identify the best performing model(s).

Four Models:

1. Model 1: Phenotype Model was defined by Simplified Severity Scale (0–4)

$$\text{Progression} = \text{intercept} + \text{baseline score} + \text{indicator}$$
2. Model 2: Genotype Model was defined by 14 disease associated SNPs

$$\text{Progression} = \text{intercept} + \text{marker1} + \text{marker 2} \dots + \text{marker14} + \text{indicator}$$
3. Model 3: Combined Model was defined by Simplified Severity Scale plus Genotype

$$\text{Progression} = \text{intercept} + \text{baseline score} + \text{marker1} + \text{marker 2} \dots + \text{marker14} + \text{indicator}$$
4. Model 4: Full Model was defined by Simplified Severity Scale plus Genotype plus Demographic and Environmental Factors. This model was derived from the full model below by backwards stepwise regression.

$$\text{Progression} = \text{intercept} + \text{marker1} + \dots + \text{marker14} + \text{baseline score} + \text{indicator} + \text{smoking} + \text{BMI} + \text{age} + \text{sex} + \text{treatment} + \text{education}$$

For Model 4, 270 missing values for smoking status were imputed to the mode (never smoked), 271 missing values for BMI were imputed to the median (26.7), and one missing value for education was imputed to the mode (more than high school) according to standard practice in order to make maximum use of the available data.

Genotype-by-treatment interactions were not investigated as estimates were not expected to be informative due to the number of years between the completion of the clinical phase of the AREDS and the ten-year follow-up. Additive coding of genotype data was applied. Specifically, genotypes were translated into the count of the number of minor alleles in the genotype (0/1/2). Baseline grade, smoking, sex, treatment (AREDS vitamin-mineral treatment assignment), and education were coded as factors; BMI and age were quantitative variables. Baseline AMD was assigned according to AREDS simplified severity scale.

Model Validation

The performance of the best performing models identified in the Training Set was assessed in the Validation Set. In time-to-event analysis, a pair of patients is called concordant if the risk of the event predicted by a model is lower for the patient who experiences the event at a later time-point. The concordance probability (C-index) is the frequency of concordant pairs among all pairs of subjects in the data. It is analogous to the area under the curve (AUC) for binary traits, and was calculated in order to measure and compare the discriminative power of the risk prediction models. All statistical analyses were performed using R version 2.13.0 (R Foundation for Statistical Computing, Vienna, Austria).

Results

The Incorporation of Genotype Improves Prediction of CNV

The univariate results for predicting CNV when time-to-event is taken into account are depicted in Table 5. Estimates for simplified severity scale are given relative to Grade 0; estimates of Treatment are given relative to Treatment Category 1 (Placebo); estimates for Smoking are given relative to 'Never Smoked.' The latter three may be assumed to have Hazard Ratio=1. The most highly associated genetic markers were SNP variants in *CFH*, *ARMS2*, *CFHR4*, *CFHR5* and Complement Component 3 (*C3*).

The estimated prediction error was similar for three of the four models. Model 1 had a Brier Score=0.099; Models 3 and 4 both had estimated Brier Scores=0.100; Model 2 had a Brier

Score=0.122. Bootstrapping, a repeated re-sampling approach, was applied to estimate confidence intervals for the Brier Scores, to determine whether the scores truly differed. All of the confidence intervals overlapped; however, the point estimate for Model 2 lay outside of (above) the confidence intervals for the other three. On this basis, Models 1, 3 and 4 were selected as the best performing models and were subjected to model validation to evaluate the clinical performance of the model in an independent set of subjects (validation set). The C-index was evaluated as a measure of clinical performance that would be anticipated in the clinical setting. Additional *ad hoc* analysis conducted to explore the incremental contribution of age and smoking (the two demographic/environmental factors that remained significant when added to the Phenotype + Genotype combined model) is shown in Table 6. These results provide an estimate of the added contribution of age and smoking to the different models.

Results of the predictions for the top performing models conducted with the validation set are shown in Table 6. Models 3 and 4 exhibited nearly equal performances and both had a higher C-index in the validation set compared to Model 1 (Phenotype). The difference in C-statistic between risk assessment based on disease at baseline (Phenotype Only =0.89; CI 0.81-0.98) and the performance achieved using a risk assessment that combined genotype with phenotype (Phenotype + Genotype =0.96 (CI 0.92–0.99) was statistically significant (p-value <0.01).

CNV Model 4 with the lowest Validation Brier score is shown in Table 7 (detail for CNV Models 1, 2 and 3 can be found in Tables 8–10, available at <http://aaojournal.org>).

The estimates of prediction error of the validation data, across time, for the three models are shown in Figure 1 (available at <http://aaojournal.org>). Models 3 and 4 show improved performances over Model 1, showing that the added contribution of genotype lowers the misclassification of progressors and non-progressors, compared to prediction based on phenotype alone.

The risk of CNV estimated over a 12 year follow-up period using Models 1, 3 and 4, respectively, are compared in Figures 2–4. Subjects are grouped by baseline disease assignment (simplified severity SCALE=0-SCALE=4) to show the risk predicted across the different stages of disease. The risk estimates achieved using Model 1 (Phenotype) for all subjects grouped by simplified severity scale assignment with or without CNV in the non-study eye are shown in Figure 2. The risk estimates for the two phenotypes (with or without CNV in fellow eye) present in subjects with SCALE=2, 3, 4 are plotted at 6 month intervals generating the four data points in each year interval. The additional stratification observed when genotype is combined with phenotype, as compared to the distribution provided under Model 1 (Phenotype) for the same set of subjects, is represented in Figure 3.

The incremental impact of comparing Model 3 (Figure 3) to Model 4 (Figure 4) with the inclusion of smoking status and age is more subtle compared to the more significant gains in granularity observed when comparing Model 1 (Figure 2) to Model 3 (Figure 3) with the addition of genotype to phenotype.

The spectrum of risk observed in AREDS subjects assigned a baseline simplified severity scale grade 3 or 4 at baseline is shown in Figure 5. Risk profiles were generated using Model 4 (Simplified Severity Scale + Genotype + Age + Smoking) to show the enhanced stratification contributed primarily by genetic predisposition. The dark middle line in each plot serves as a reference for the risk estimation based on phenotype alone. The three risk groups; high, moderate and low represent the top, middle and bottom 30% of the patient population at each of the two grades (3, 4). The outliers (< 5% and > 95%) are represented by dashed-line boundaries reflecting the most extreme patients in each category.

The added predictive gains of combining genetics with phenotypic assessment to assess the risk of progressing to CNV at 2, 5 and 10 years are shown in Table 11. Risk estimates are provided for AREDS subjects assigned simplified severity scale 3 (upper panel) or 4 (lower panel) at baseline. The estimated risk listed as 50% reflects the median risk of all subjects in the group reflective of the reported risk if phenotype alone was used to estimate risk of CNV. Without genotype included, all subjects designated SCALE=4 would be advised they had a 68% risk of conversion to CNV in 10 years based on phenotype alone. With genotype, age and smoking status included, patients with the same SCALE=4 can be stratified into three categories; the lowest genetic burden would have a 30% probability of conversion to CNV, while the highest group of patients would have an 90% probability of conversion (see bold values highlighted in Table 11). Of the 238 Validation Set subjects with the most clinically advanced phenotypes (SCALE=2, 3, 4), 15% of the subjects were identified as having a probability of conversion to CNV over 75%. The identification of these high risk subjects would be masked within the SCALE 4 group if risk assessment was conducted using phenotype alone. Without genotyping, all SCALE 4 subjects would be assigned a single risk estimate based on the mean risk of conversion for the entire group. The benefit of including genotype is in providing a personalized risk assessment for each patient instead of applying predictions based on the average risk of the population presenting with the same stage of disease.

Risk Score as Measure of Genetic Load

The additional stratification provided by incorporating genetics with baseline disease assessment (AREDS Simplified Severity Scale) is depicted in Figure 6 (available at <http://aaojournal.org>). The boxplots reveal the spectrum of genetic risk within a group of subjects classified according to simplified severity scale 0 (no disease) through 4. A strong concordance is observed between the median probability of lifetime risk of CNV for each simplified severity scale category group 0–4 (shown as the dark solid line in Figure 6, available at <http://aaojournal.org>) indicating a strong correlation between genetic predisposition and baseline disease severity.

Prediction of GA

The univariate results for GA when time-to-event is taken into account are depicted in Table 12. Estimates for simplified severity scale are given relative to Grade 0; estimates of Treatment are given relative to Treatment Category 1 (Placebo); estimates for Smoking are given relative to 'Never Smoked.' The latter three may be assumed to have Hazard Ratio=1. The impact of genetic markers appears less pronounced compared to the CNV model, although they remain integral to the accuracy of prediction. The SNPs with the most significant impact remain those detecting variants in *CFH*, *ARMS2*, *CFHR4*, *CFHR5* and *C3*.

Evaluation of the four training models for GA prediction was conducted in the same manner as was performed for the CNV endpoint. The Brier Score was lowest for GA Model 3 (Brier Score=0.072) compared to the other models that had scores 0.079, 0.084 and 0.074 for Models 1, 2 and 4, respectively. Details for GA Models 1–4 are shown in Tables 13–16 (available at <http://aaojournal.org>).

Predictions for Model 3 (Simplified Severity Scale + Genotype) were obtained using the validation set generating a C-Index of 0.96 and a Brier Score of 0.058.

Figures 7 and 8 show the risk of GA estimated over a 12 year follow-up period using Models 1 and 3, respectively. Subjects are grouped by baseline disease assignment (simplified severity SCALE=0-SCALE=4) to show the risk predicted across the different

stages of disease. Figure 7 shows the risk estimates achieved using Model 1 (Phenotype) for all subjects grouped by simplified severity scale assignment with or without GA in the non-study eye. The risk estimates for the two phenotypes are plotted at semi-annual and annual time points generating the four data points in each year interval. While not statistically significant, the contribution of advanced disease (GA) in the non-study fellow eye appears more pronounced in GA risk prediction compared to the impact of advanced disease (CNV) in the non-study eye for CNV prediction (Figure 2).

Figure 8 shows the range of predictions observed with Model 3. The added granularity of genetics in predicting the onset of GA is similar to that observed in the CNV progression model. Genetics further stratifies patients presenting with the same baseline stage of disease. Figure 9 (available at <http://aaojournal.org>) shows the estimates of prediction error over time for Model 3, as compared to the Kaplan Meier curve reflecting the GA events recorded across the follow-up period.

Follow-up Analyses for GA Model 3

A more parsimonious model was derived from GA Model 3 by applying backwards stepwise regression to the training data. The purpose of the backward regression was to remove the factors that did not contribute to the overall performance of the model in an effort to arrive at a more efficient panel. The resultant model is shown in Table 17. This model had a Brier Score=0.074, slightly weaker than before the backwards regression was applied (Brier Score=0.072).

The reduced GA model was assessed using the validation data and the results shown in Table 18 (available at <http://aaojournal.org>) were obtained. Performance was slightly compromised in the reduced model (Table 18), as shown in the drop in C-index from 0.96 to 0.94, but reflects a more efficient design. Interestingly, several of the genetic variants appeared to have a lower impact on the progression to GA compared to the influence observed in the progression to CNV. Notably, only one SNP remained in the final GA progression model, a variant in the *CFH-CFHR-F13B* gene cluster on chromosome 1, compared to eight SNPs that remained clinically relevant in the CNV progression model. The additional SNPs representing variants in *CFHR4*, *CFHR5* and *F13B* are more influential in the progression to CNV than GA.

Discussion

This study was conducted to evaluate four predictive models for separately estimating the conversion of early stage AMD to the more advanced forms of the disease, specifically, CNV and GA. The objective was to determine if risk prediction, based purely on baseline disease assignment, was as accurate as an estimation that also included a patient's genetic predisposition (understanding that some DNA markers are protective while some are disease-contributing), using clinically relevant DNA markers.⁹ The results demonstrated that the highest performing models for separate prediction of conversion to CNV and GA contained parameters that captured both baseline disease and genetic variation. Without including genetic predisposition into patient risk assessment, one cannot objectively capture the variability observed across a group of patients using an assessment based on disease stage alone. The difference in C-statistic obtained in the CNV risk assessment based on disease at baseline (Phenotype Only) was statistically significant (p-value <0.01) when compared to the performance achieved using a risk assessment that combined genotype with phenotype. Providing estimates of disease progression without genetic information may result in over-estimating risk in patients with low risk genotypes and under-estimating risk in patients with high risk genetic profiles; the result of which may lead to sub-optimal patient management.

Previous reports where only one or two SNPs within the complex genetic structure of the CFH gene were studied lack the resolution provided by the eight CFH/CFHR SNPs used in the current model to distinguish clinically relevant haplotypes combinations. The powerful C-statistic achieved in the present predictive model is also likely due to use of a single phenotypic CNV endpoint instead of the performance metrics reported in earlier models that combined prediction of both advanced forms of late-stage disease (GA and CNV). Combining the prediction of two phenotypically distinct endpoints that appear to be differentially influenced by several SNPs and smoking history might compromise the overall accuracy of prediction compared to the accuracy obtained in a model predicting CNV alone.

Analysis of the 2,415 AREDS study subjects, stratified according to AREDS simplified severity scale, revealed the benefits of assessing genetic factors in combination with patient disease stage. The genetic heterogeneity observed was not limited to a single baseline grade, emphasizing the enhanced stratification genetic information provides across early, intermediate and late stages of disease. Development of a test that provides physicians with the genotypes of clinically validated genetic variants combined with simple phenotypic assignment to predict time to advanced disease onset, specifically, at 2, 5, and 10 years, provides a more robust clinical tool to anticipate a patient's individual progression profile, and provide an opportunity to improve patient management. If a physician wished to decide which of two arbitrary patients was at higher risk, the number of incorrect assignments would be reduced by 66% using a model that predicts progression based on a patient's phenotype, genotype, age and smoking status compared to an estimate based on phenotype alone. This difference represents a potential improvement in treatment for millions of patients with AMD. An improvement in CNV risk prediction is an important factor in clinical management of AMD patients by medical professionals. As for improved outcomes, investment in prospective studies will be required to generate data that can be used to assess this question.

We have also revealed that the impact of the disease-associated genetic markers was not equally influential in progression to CNV compared to progression to GA, with specific markers, (*e.g.*, ARMS2), showing a more influential contribution in the CNV model. Based on this observation, originally reported by Sobrin¹⁰ and replicated in this study, separate independent models should be used to predict progression to CNV and GA to account for parameters, including genetic markers and smoking status, that do not appear to influence both endpoints equally. Prior published reports of predictive models for both CNV and GA based on the shared endpoint of a blinding outcome may be misrepresentative from a biological standpoint if the phenotypes of advanced disease are actually based on different biological mechanisms.^{7,8} One limitation of the predictive model for GA in this study was that the endpoint was limited to the vision-threatening form of central GA rather than all GA. By including these values, the model we present here might be more biologically relevant.

The results presented in this study also highlight the value of using the statistically robust time-to-onset methodology, which exploits the full longitudinal data available in the AREDS database, as opposed to cross-sectional analysis. Developing predictive models using this methodology incorporates the variable of time to disease onset, to enhance accuracy of prediction, as opposed to cross sectional studies limited to a single disease state endpoint. Given the variability in time points at which patients converted to CNV or GA during the ten year follow-up period, this analysis strategy maximizes the information available while also accounting for the often transient presentation of CNV that might otherwise be missed in a single time point evaluation. The use of annual visit data to identify a "progressor" in lieu of final subject summary assignment based on re-grading and final status review might reflect a less stringent definition of advanced disease than evaluated in

earlier studies and might result in slightly higher reported progression rates in this study. The utilization of interim data was believed to capture a more accurate reflection of the assignments anticipated in the clinical setting.

Recently, the American Academy of Ophthalmology (AAO) has recommended against the use of routine genetic screening in genetically complex disorders like age-related macular degeneration.²² Absent a therapeutic to delay progression in the early stages of AMD, the utility of genetic screening in asymptomatic individuals is of limited utility. However, directed testing in patients to identify those individuals that have a high genetic burden combined with symptoms of progressive disease will allow physicians to concentrate care in a subset of the AMD patient population with the highest risk of losing vision.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Financial Support: Sequenom Center for Molecular Medicine (SCMM), San Diego, CA 92121. The sponsor participated in the design of the study, conducting the study, data collection, data management, data analysis, interpretation of the data, manuscript preparation and review.

This study was conducted in compliance with the Coriell Cell Repositories Institutional Review Board, in accordance with Department of Health and Human Services (45 CFR Part 46). The dataset used for the analysis contained in this manuscript was obtained from the National Eye Institute-Age-Related Eye Disease Study (NEI-AREDS) Genetic Repository. Funding support for AREDS was provided by the National Eye Institute grant N01-EY-0-2127, National Institutes of Health, Bethesda, Maryland.

References

1. National Eye Institute. [Accessed February 4, 2013] Prevalence of Age-Related Macular Degeneration in the United States [NEI Statistics and Data]. Available at: <http://www.nei.nih.gov/eyedata/pbd4.asp>.
2. Tomany SC, Wang JJ, van Leeuwen R, et al. Risk factors for incident age-related macular degeneration: pooled findings from 3 continents. *Ophthalmology*. 2004; 111:1280–1287. [PubMed: 15234127]
3. AMD Alliance International. Basic facts about AMD. Available at: http://www.amdalliance.org/information_overview_basic_facts.html.
4. Lim JH, Wickremasinghe SS, Xie J, et al. Delay to treatment and visual outcomes in patients treated with anti-vascular endothelial growth factor for age-related macular degeneration. *Am J Ophthalmol*. 2012; 153:678–686. [PubMed: 22245460]
5. Muether PS, Hoerster R, Hermann MM, et al. Long-term effects of ranibizumab treatment delay in neovascular age-related macular degeneration. *Graefes Arch Clin Exp Ophthalmol*. In press.
6. Muether PS, Hermann MM, Koch K, Fauser S. Delay between medical indication to anti-VEGF treatment in age-related macular degeneration can result in a loss of visual acuity. *Graefes Arch Clin Exp Ophthalmol*. 2011; 249:633–637. [PubMed: 20865421]
7. Seddon JM, Reynolds R, Yu Y, et al. Risk models for progression to advanced age-related macular degeneration using demographic, environmental, genetic, and ocular factors. *Ophthalmology*. 2011; 118:2203–2211. [PubMed: 21959373]
8. Klein ML, Francis PJ, Ferris FL III, et al. Risk assessment model for development of advanced age-related macular degeneration. *Arch Ophthalmol*. 2011; 129:1543–1550. [PubMed: 21825180]
9. Hageman GS, Gehrs K, Lejnine S, et al. Clinical validation of a genetic model to estimate the risk of developing choroidal neovascular age-related macular degeneration. *Hum Genomics*. 2011; 5:420–440. [PubMed: 21807600]

10. McCarthy LC, Newcombe PJ, Whittaker JC, et al. Predictive models of choroidal neovascularization and geographic atrophy incidence applied to clinical trial design. *Am J Ophthalmol*. 2012; 154:568–578. [PubMed: 22704140]
11. Grassmann F, Fritsche LG, Keilhauer CN, et al. Modelling the genetic risk in age-related macular degeneration. *PLoS ONE* [serial online]. 2012; 7:e37979. Available at: <http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0037979>.
12. Yu Y, Reynolds R, Rosner B, et al. Prospective assessment of genetic effects on progression to different stages of age-related macular degeneration using multistate Markov models. *Invest Ophthalmol Vis Sci*. 2012; 53:1548–1556. [PubMed: 22247473]
13. Spencer KL, Olson LM, Schnetz-Boutaud N, et al. Using genetic variation and environmental risk factor data to identify individuals at high risk for age-related macular degeneration. *PLoS One* [serial online]. 2011; 6:e17784. Available at: <http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0017784>.
14. Chen Y, Zeng J, Zhao C, et al. Assessing susceptibility to age-related macular degeneration with genetic markers and environmental factors. *Arch Ophthalmol*. 2011; 129:344–351. [PubMed: 21402993]
15. Seddon JM, Reynolds R, Rosner B. Associations of smoking, body mass index, dietary lutein, and the LIPC genetic variant rs10468017 with advanced age-related macular degeneration. *Mol Vis* [serial online]. 2010; 16:2412–2424. Available at: <http://www.molvis.org/molvis/v16/a259/>.
16. Hughes AE, Orr N, Patterson C, et al. Neovascular age-related macular degeneration risk based on CFH, LOC387715/HTRA1, and smoking. *PLoS Med* [serial online]. 2007; 4:e355. Available at: <http://www.plosmedicine.org/article/metrics/info%3Adoi%2F10.1371%2Fjournal.pmed.0040355>.
17. Hughes AE, Orr N, Esfandiary H, et al. A common CFH haplotype, with deletion of CFHR1 and CFHR3, is associated with lower risk of age-related macular degeneration. *Nat Genet*. 2006; 38:1173–1177. [PubMed: 16998489]
18. Hageman GS, Anderson DH, Johnson LV, et al. A common haplotype in the complement regulatory gene factor H (HF1/CFH) predisposes individuals to age-related macular degeneration. *Proc Natl Acad Sci U S A*. 2005; 102:7227–7232. [PubMed: 15870199]
19. Sobrin L, Ripke S, Yu Y, et al. Heritability and genome-wide association study to assess genetic differences between advanced age-related macular degeneration subtypes. *Ophthalmology*. 2012; 119:1874–1885. [PubMed: 22705344]
20. Yanagisawa S, Kondo N, Miki A, et al. Difference between age-related macular degeneration and polypoidal choroidal vasculopathy in the hereditary contribution of the A69S variant of the age-related maculopathy susceptibility 2 gene (ARMS2). *Mol Vis* [serial online]. 2011; 17:3574–3582. Available at: <http://www.molvis.org/molvis/v17/a384/>.
21. Age-Related Eye Disease Study Research Group. A simplified severity scale for age-related macular degeneration: AREDS report no. 18. *Arch Ophthalmol*. 2005; 123:1570–1574. [PubMed: 16286620]
22. Stone EM, Aldave AJ, Drack AV, et al. Recommendations for genetic testing of inherited eye diseases: report of the American Academy of Ophthalmology Task Force on Genetic Testing. *Ophthalmology*. 2012; 119:2408–2410. [PubMed: 22944025]
23. Age-Related Eye Disease Study Research Group. The Age-Related Eye Disease Study (AREDS): design implications. AREDS report no. 1. *Control Clin Trials*. 1999; 20:573–600. [PubMed: 10588299]
24. Age-Related Eye Disease Study Research Group. The Age-Related Eye Disease Study severity scale for age-related macular degeneration: AREDS report no. 17. *Arch Ophthalmol*. 2005; 123:1484–1498. [PubMed: 16286610]

Figure 2. Choroidal neovascularization (CNV) risk estimation of 824 validation set subjects grouped by simplified severity scale based on Model 1 (Phenotype). For each semi-annual time point, the green line (1) corresponds to patients with CNV in the fellow eye and the red line (0) reflects subjects without unilateral CNV at baseline.

Figure 3.

Choroidal neovascularization (CNV) risk estimations of 824 validation set subjects grouped by simplified severity scale based on Model 3 (Phenotype + Genotype) shows the added granularity contributed by genetic information. The red dotted line represents a quantile regression line at 75% probability of risk of CNV while the blue dotted line represents a quantile regression line at 25% probability of CNV.

Figure 4.

Choroidal neovascularization (CNV) risk estimation of 824 validation set subjects grouped by simplified severity scale based on Model 4 (Phenotype + Genotype + Age + Smoking Status) shows the incremental change from Model 3 (Figure 3) with the addition of age and smoking status. The red dotted line represents a quantile regression line at 75% probability of risk of CNV while the blue dotted line represents a quantile regression line at 25% probability of CNV.

Figure 5.

Risk progression profiles based on Model 4 (Simplified Severity Scale + Genotype + Age + Smoking) of 824 validation set subjects assigned simplified severity scale grades 3 (SCALE=3) and 4 (SCALE=4). Plots reveal the spectrum of risk contributed primarily by genetic predisposition. The dark middle line in each plot serves as a reference for the risk estimation based on phenotype alone (simplified severity scale). The outliers (< 5% and > 95%) are represented by dashed-lines reflecting the most extreme patients in each category. CNV = choroidal neovascularization.

Figure 7.

Geographic atrophy (GA) risk estimation of 824 subjects in validation set grouped by simplified severity scale based on Model 1 (Phenotype). For each semi-annual time point, the green line (1) corresponds to patients with GA in the fellow eye and the red line (0) reflects subjects without unilateral GA at baseline. The impact of advanced disease (GA) in the non-study eye at baseline is reflected in the steeper risk progression line observed in subjects with SCALE=2, 3 4 compared to the lower line, reflecting risk of subjects without unilateral GA at baseline.

Figure 8.

Geographic atrophy (GA) risk estimation of 824 validation set subjects grouped by simplified severity scale based on Model 3 (Simplified Severity Scale + Genotype) shows the added granularity contributed by genetics. The red dotted line represents a quantile regression line at 75% probability of risk of choroidal neovascularization (CNV) while the blue dotted line represents a quantile regression line at 25% probability of CNV.

Table 1

Genetic Markers

DNA Single Nucleotide Polymorphism (SNP)	Chromosome	Gene
rs1061170	1	Complement Factor H (<i>CFH</i>) (exon 9)
rs2274700	1	Complement Factor H (<i>CFH</i>) (exon 10)
rs403846	1	Complement Factor H (<i>CFH</i>) (intron 14)
rs12144939	1	Complement Factor H (<i>CFH</i>) (intron 15)
rs1409153	1	Complement Factor H-Related 4 (<i>CFHR4</i>)
rs1750311	1	Complement Factor H-Related 5 (<i>CFHR5</i>)
rs10922153	1	Complement Factor H-Related 5 (<i>CFHR5</i>)
rs698859	1	Coagulation factor XIII B subunit (<i>F13B</i>)
rs2990510	1	Coagulation factor XIII B subunit (<i>F13B</i>)
rs9332739	6	Complement Component 2 (<i>C2</i>)
rs641153	6	Complement Factor B (<i>CFB</i>)
rs10490924	10	Age-related maculopathy susceptibility protein 2 (<i>ARMS2</i>)
rs11200638	10	High temperature requirement factor A1 (<i>HTRA1</i>)
rs2230199	19	Complement Component 3 (<i>C3</i>)

Table 2

Training and Validation Sets

Measure	Training Set	Validation Set
Count	1589	826
Percent Progression to CNV	25	24
Percent Progression to GA	16	16
Percent All Advanced AMD (GA and/or CNV)	30	29
Age (Years; Mean)	68.75	68.23
Percent Female	58	57
Percent Current Smoker	5	5

CNV = choroidal neovascularization; GA = geographic atrophy; AMD = age-related macular degeneration

Table 3

Demographic Characteristics

Measure	CNV Onset	GA Onset	All
Count	603	379	2415
Baseline Score	0: 23 1: 44 2: 117 3: 160 4: 256 NA: 3	0: 11 1: 36 2: 58 3: 95 4: 177 NA: 2	0: 940 1: 417 2: 397 3: 287 4: 368 NA: 6
Smoking Status	Never: 220 Past: 276 Current: 44 NA: 63	Never: 161 Past: 157 Current: 23 NA: 38	Never: 1028 Past: 997 Current: 120 NA: 270
BMI (Mean (SE); NA count)	27.93 (0.18) NA: 63	27.99 (0.24) NA: 38	27.42 (0.10) NA: 271
Age (Mean (SE); NA count)	70.40 (0.21) NA: 0	69.71 (0.27) NA: 0	68.57 (0.10) NA: 0
Sex	Female: 359 Male: 244 NA: 0	Female: 223 Male: 156 NA: 0	Female: 1393 Male: 1022 NA: 0
Treatment Category	1: 138 2: 155 3: 168 4: 142 NA: 0	1: 75 2: 92 3: 107 4: 105 NA: 0	1: 720 2: 770 3: 466 4: 459 NA: 0
High School (HS) Education	HS: 239 >HS: 364 NA: 0	HS: 150 >HS: 229 NA: 0	HS: 784 >HS: 1630 NA: 1

CNV = choroidal neovascularization; GA = geographic atrophy; BMI = body mass index; NA = data not available; SE = standard error

Table 5

CNV Univariate Analysis Using a Cox Proportional Hazards Model

Variable	Hazard Ratio	ln (SE)	95% Confidence Interval	p-value
Simplified Severity Scale Grade 4	58.59	0.2989	32.62,105.25	<2.00E-16
Simplified Severity Scale Grade 3	37.70	0.3047	20.75,68.5	<2.00E-16
Simplified Severity Scale Grade 2	16.96	0.3099	9.24,31.14	<2.00E-16
CNV in Non-study Eye at Baseline	3.94	0.1055	3.21,4.85	<2.00E-16
SNP rs10490924 (ARMS2)	2.20	0.0686	1.93,2.52	<2.00E-16
SNP rs11200638 (HTRA1)	2.20	0.0687	1.93,2.52	<2.00E-16
SNP rs1409153 (CFHR4)	2.05	0.0711	1.79,2.36	<2.00E-16
SNP rs1061170 (CFH)	1.94	0.0713	1.69,2.23	<2.00E-16
SNP rs403846 (CFH)	0.48	0.0737	0.41,0.55	<2.00E-16
SNP rs2274700 (CFH)	0.39	0.0926	0.33,0.47	<2.00E-16
SNP rs10922153 (CFHR5)	0.53	0.0768	0.46,0.62	1.11E-16
Age	1.09	0.0106	1.07,1.11	2.22E-15
SNP rs12144939 (CFH)	0.43	0.1289	0.33,0.55	5.48E-11
SNP rs1750311 (CFHR5)	0.65	0.0852	0.55,0.77	3.57E-07
Simplified Severity Scale Grade 1	5.34	0.3433	2.73,10.47	1.05E-06
AREDS Treatment: Zinc Supplement	1.95	0.1406	1.48,2.57	2.03E-06
SNP rs2990510 (F13B)	1.34	0.0715	1.17,1.55	3.76E-05
SNP rs2230199 (C3)	1.38	0.0779	1.18,1.60	4.23E-05
SNP rs641153 (Factor B)	0.52	0.1773	0.37,0.74	0.0003
Current Smoker	2.00	0.1999	1.35,2.96	0.0005
AREDS Treatment: Antioxidants + Zinc	1.63	0.1460	1.23,2.17	0.0008
Education Level at or Above High School	0.71	0.1021	0.58,0.87	0.0008
SNP rs9332739 (CFH)	0.46	0.2597	0.28,0.76	0.0028
Previous Smoker (6 Months)	1.36	0.1033	1.11,1.67	0.0028
Body Mass Index (BMI)	1.03	0.0098	1.01,1.05	0.0045
SNP rs698859 (F13B)	1.10	0.0708	0.96,1.27	0.1621
Sex	0.94	0.1012	0.77,1.15	0.5447
AREDS Treatment: Antioxidant Supplement	0.96	0.1442	0.73,1.28	0.7904

CNV = choroidal neovascularization; ln (SE) = natural log of the standard error; SNP = single nucleotide polymorphism; AREDS = Age-Related Eye Disease Study

Table 6

Validation Estimates for Time-To-Event Analysis of CNV

Best Performing Models	C-index	Brier Score
Model 1: Phenotype: Simplified Severity Scale + Assessment of CNV in fellow eye	0.89	0.073
Model 3: Combined: Phenotype + Genotype	0.96	0.070
Model 4: Full: Phenotype + Genotype + Demographic + Environmental	0.96	0.069
Models Evaluated in <i>ad hoc</i> Analysis to Explore Contribution of Age + Smoking		
Model 2: Genotype: 14 Disease associated SNPs (evaluated in primary and <i>ad hoc</i> analysis)	0.84	0.090
Model 5: Smoking + Age Only	0.63	0.099
Model 6: Genotype + Smoking + Age	0.86	0.087
Model 7: Phenotype + Smoking + Age	0.90	0.072

CNV = choroidal neovascularization

Table 7

CNV Training Set Model 4: Full Model-Simplified Severity Scale plus Genetics plus Environment (Brier Score=0.100)

Variable	Hazard Ratio	ln (SE)	95% Confidence Interval	p-value
Simplified Severity Scale 1	4.76	0.34	2.43, 9.34	5.61E-06
Simplified Severity Scale 2	12.66	0.31	6.87, 23.36	4.44E-16
Simplified Severity Scale 3	26.56	0.31	14.53, 48.58	<2.00E-16
Simplified Severity Scale 4	35.89	0.30	19.75, 65.21	<2.00E-16
SNP rs1061170	0.75	0.13	0.58, 0.97	0.0294
SNP rs2274700	0.77	0.16	0.56, 1.06	0.1052
SNP rs12144939	0.73	0.21	0.48, 1.11	0.1401
SNP rs1409153	1.55	0.15	1.15, 2.09	0.0045
SNP rs1750311	1.50	0.24	0.93, 2.43	0.0958
SNP rs10922153	0.63	0.29	0.36, 1.11	0.1119
SNP rs698859	0.67	0.21	0.44, 1.01	0.0554
SNP rs2990510	0.66	0.23	0.42, 1.04	0.0705
SNP rs10490924	1.46	0.07	1.27, 1.67	1.43E-07
SNP rs2230199	1.12	0.08	0.96, 1.31	0.1362
SNP rs9332739	0.57	0.26	0.34, 0.96	0.0351
SNP rs641153	0.65	0.18	0.46, 0.93	0.0169
Previous Smoker (6 Months)	1.24	0.11	1.01, 1.52	0.0438
Current Smoker	1.78	0.21	1.18, 2.67	0.0056
Age	1.05	0.01	1.03, 1.07	5.81E-06

CNV = choroidal neovascularization; ln (SE) = natural log of the standard error; SNP = single nucleotide polymorphism

Table 11

Risk Estimates, By Specified Percentile, For AREDS Subjects Assigned Simplified Severity Scale 3 (Upper Panel) or 4 (Lower Panel) At Baseline. Tables Reveal The Spectrum Of Risk Values Contributed Primarily By Genetic Predisposition.

Risk Estimates for Simplified Severity Scale 3			
	Year 2	Year 5	Year 10
Very Low (5%)	4.1%	9.8%	22.3%
Low (20%)	6.9%	16.3%	35.3%
Moderate (50%)	12.0%	27.2%	53.9%
High (80%)	16.1%	35.4%	65.6%
Very High (95%)	23.0%	47.8%	79.6%
Risk Estimates for Simplified Severity Scale 4			
	Year 2	Year 5	Year 10
Very Low (5%)	5.6%	13.4%	29.6%
Low (20%)	10.4%	23.8%	48.6%
Moderate (50%)	16.9%	36.9%	67.6%
High (80%)	23.3%	48.3%	80.0%
Very High (95%)	30.8%	59.9%	89.3%

AREDS = Age-Related Eye Disease Study

Table 12

Univariate Analysis for GA using a Cox Proportional Hazards Model

Variable	Hazard Ratio	ln (SE)	95% Confidence Interval	p-value
Simplified Severity Scale Grade 4	53.78	0.3893	25.07,115.35	<2.00E-16
Simplified Severity Scale Grade 3	31.09	0.3988	14.23,67.94	<2.00E-16
SNP rs10490924 (ARMS2)	2.03	0.0870	1.71,2.41	3.33E-16
SNP rs11200638 (HTRA1)	2.02	0.0867	1.70,2.39	5.55E-16
SNP rs2274700 (CFH)	0.41	0.1178	0.33,0.52	6.92E-14
SNP rs403846 (CFH)	0.50	0.0939	0.42,0.60	2.12E-13
SNP rs1061170 (CFH)	1.96	0.0917	1.63,2.34	2.48E-13
SNP rs1409153 (CFHR4)	1.88	0.0904	1.58,2.25	2.62E-12
Simplified Severity Scale Grade 2	11.93	0.4131	5.31,26.82	1.95E-09
SNP rs10922153 (CFHR5)	0.58	0.0970	0.48,0.71	2.98E-08
GA in Non-study Eye at Baseline	3.75	0.2389	2.35,5.99	3.13E-08
Simplified Severity Scale Grade 1	7.85	0.4276	3.40,18.15	1.44E-06
SNP rs12144939 (CFH)	0.46	0.1619	0.33,0.63	1.63E-06
AREDS Treatment: Antioxidants + Zinc	2.47	0.1906	1.70,3.59	2.05E-06
AREDS Treatment: Zinc Supplement	2.44	0.1906	1.68,3.54	2.97E-06
Age	1.05	0.0134	1.03,1.08	0.0001
SNP rs2230199 (C3)	1.41	0.0993	1.16,1.71	0.0005
SNP rs1750311 (CFHR5)	0.72	0.1061	0.59,0.89	0.0021
Education Level at or Above High School	0.69	0.1298	0.54,0.90	0.0049
Body Mass Index (BMI)	1.03	0.0126	1.01,1.06	0.0106
SNP rs2990510 (F13B)	1.26	0.0919	1.05,1.51	0.0122
SNP rs641153 (Factor B)	0.59	0.2175	0.38,0.90	0.0151
SNP rs9332739 (CFH)	0.42	0.3551	0.21,0.85	0.0154
SNP rs698859 (F13B)	1.16	0.0900	0.97,1.38	0.0998
Current Smoker	1.35	0.2732	0.79,2.30	0.2763
AREDS Treatment: Antioxidant Supplement	1.18	0.1965	0.80,1.74	0.3927
Previous Smoker (6 Months)	1.08	0.1315	0.83,1.40	0.5618
Sex	0.97	0.1291	0.75,1.25	0.8119

GA = geographic atrophy; ln (SE) = natural log of the standard error; SNP = single nucleotide polymorphism; AREDS = Age-Related Eye Disease Study

Table 17

GA Combined Model 3: Simplified Severity Scale plus Genotype after Backwards Stepwise Regression (Brier Score=0.074)

Variable	Hazard Ratio	ln (SE)	95% Confidence Interval	p-value
Simplified Severity Scale 1	6.97	0.4284	3.01, 16.14	5.87E-06
Simplified Severity Scale 2	9.33	0.4153	4.13, 21.05	7.55E-08
Simplified Severity Scale 3	23.29	0.4022	10.59, 51.22	5.00E-15
Simplified Severity Scale 4	34.81	0.3961	16.02, 75.65	<2.00E-16
SNP rs2274700 (<i>CFH</i>)	0.64	0.1194	0.51, 0.81	0.0002
SNP rs10490924 (<i>ARMS2</i>)	1.35	0.0895	1.14, 1.61	0.0007
SNP rs2230199 (<i>C3</i>)	1.22	0.0979	1.01, 1.48	0.0430
SNP rs9332739 (<i>C2</i>)	0.58	0.3602	0.29, 1.18	0.1335
GA in Non-study Eye at Baseline	1.76	0.2411	1.1, 2.83	0.0189

GA = geographic atrophy; ln (SE) = natural log of the standard error; SNP = single nucleotide polymorphism