CFH and *ARMS2* genetic risk determines progression to neovascular age-related macular degeneration after antioxidant and zinc supplementation

Demetrios G. Vavvas (Δημήτριος Γ. Βάββας)[a], Kent W. Small[b], Carl C. Awh[c], Brent W. Zanke[d], Robert J. Tibshirani[e,f,1], and Rafal Kustra[g]

[a]Department of Ophthalmology Retina Service, Massachusetts Eye and Ear Institute, Harvard Medical School, Boston, MA 02114; [b]Macula and Retina Institute, Los Angeles, CA 90048; [c]Tennessee Retina, Nashville, TN 37203; [d]Department of Medical Affairs, Arctic Medical Laboratories, Grand Rapids, MI 49504; [e]Department of Biomedical Data Science, Stanford University, Stanford, CA 94305; [f]Department of Statistics, Stanford University, Stanford, CA 94305; and [g]Dalla Lana School of Public Health, University of Toronto, Toronto, ON M5T 3M7, Canada

Contributed by Robert J. Tibshirani, December 12, 2017 (sent for review October 18, 2017; reviewed by Tom Friberg and J. Sunil Rao)

We evaluated the influence of an antioxidant and zinc nutritional supplement [the Age-Related Eye Disease Study (AREDS) formulation] on delaying or preventing progression to neovascular AMD (NV) in persons with age-related macular degeneration (AMD). AREDS subjects (n = 802) with category 3 or 4 AMD at baseline who had been treated with placebo or the AREDS formulation were evaluated for differences in the risk of progression to NV as a function of *complement factor H* (*CFH*) and *age-related maculopathy susceptibility 2* (*ARMS2*) genotype groups. We used published genetic grouping: a two-SNP haplotype risk-calling algorithm to assess *CFH*, and either the single SNP rs10490924 or 372_815del443ins54 to mark *ARMS2* risk. Progression risk was determined using the Cox proportional hazard model. Genetics–treatment interaction on NV risk was assessed using a multiiterative bootstrap validation analysis. We identified strong interaction of genetics with AREDS formulation treatment on the development of NV. Individuals with high *CFH* and no *ARMS2* risk alleles and taking the AREDS formulation had increased progression to NV compared with placebo. Those with low *CFH* risk and high *ARMS2* risk had decreased progression risk. Analysis of *CFH* and *ARMS2* genotype groups from a validation dataset reinforces this conclusion. Bootstrapping analysis confirms the presence of a genetics–treatment interaction and suggests that individual treatment response to the AREDS formulation is largely determined by genetics. The AREDS formulation modifies the risk of progression to NV based on individual genetics. Its use should be based on patient-specific genotype.

ophthalmology | macular degeneration | bootstrap validation | genetic effect modification | statistical interaction

Age-related macular degeneration (AMD) is the leading cause of visual disability in the industrialized world and the third leading cause globally (1). Approximately 11 million individuals are affected with AMD in the United States, with a global prevalence of 170 million, projected to increase to 288 million by the year 2050. In 2009, the direct annual health care cost due to AMD in the US was $4.6 billion (2).

AMD preferentially damages the macula, the central region of the retina (3). AMD may be classified as early, intermediate, or advanced, based on macular phenotype and visual acuity. Blindness due to AMD is typically caused by the advanced stage of the disease, which takes two principal forms. Neovascular AMD (NV) refers to pathologic angiogenesis and its sequelae and is characterized by relatively rapid loss of central vision. Central geographic atrophy (GA) refers to localized atrophy of central macular tissue and associated structures and is characterized by a more gradual progressive loss of vision (3). The risk of developing AMD is influenced by a complex interaction of age, environment, and genetics. Genetic factors add to retinal phenotype as predictors of advanced disease (4). Polymorphisms in the *complement factor H* (*CFH*) and *age-related maculopathy susceptibility 2* (*ARMS2*) genes have the greatest impact on the progression to advanced AMD (5–7).

The Age-Related Eye Disease Study (AREDS) concluded that the AREDS formulation, a combination of high-dose antioxidants (β-carotene, vitamin C, and vitamin E) and high-dose zinc, reduced the 5-y risk of progression from intermediate to advanced age-related macular degeneration by 25% (8). Although advanced AMD was defined in the AREDS as either NV or central GA, the demonstrated reduction in progression to advanced AMD was due to a decrease in progression to NV and not by decreased progression to central GA, even after long-term evaluation (8, 9).

Multiple publications have evaluated the influence of genetic risk on the response to the AREDS formulation (10–13). These analyses are all derived from various data subsets from the AREDS, the only placebo-controlled, long-term study of the effect of nutritional therapy on AMD progression. These publications address the phenomenon of interaction, a statistical term indicating that the effect of one independent variable on a dependent variable

Significance

Age-related macular degeneration (AMD) is the leading cause of severe vision loss in the elderly and has major economic and quality-of-life impact. Prophylactic high-dose zinc and antioxidant supplements treatments are typically recommended with the assumption of homogeneously distributed benefit and risk of developing neovascular AMD. We show that individual variation at *complement factor H* and *age-related maculopathy susceptibility 2*, genes which predispose to AMD, also determines the effectiveness of nutritional prophylaxis. Some individuals paradoxically experience worsening disease with treatment, while others experience greater than average benefit. These divergent responses are difficult to identify when treatment effects have long latency. Understanding individual variations in prophylactic treatment response should inform future research and optimize health outcomes.

Author contributions: D.G.V., C.C.A., B.W.Z., and R.K. designed research; R.J.T. and R.K. performed research; R.J.T. and R.K. contributed new reagents/analytic tools; C.C.A., B.W.Z., R.J.T., and R.K. analyzed data; and D.G.V., K.W.S., C.C.A., B.W.Z., R.J.T., and R.K. wrote the paper.

Reviewers: T.F., University of Pittsburgh; and J.S.R., University of Miami.

Conflict of interest statement: B.W.Z. is the director of Arctic Medical Laboratories and founder and an equity holder of ArcticDx, Inc. (>5%), which owns patents relevant to the results; C.C.A. is a medical consultant and equity holder of ArcticAx Inc. (<1%); and R.K. is a technical consultant and equity holder of ArcticAx Inc. (<1%).

This open access article is distributed under Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND).

[1]To whom correspondence should be addressed. Email: tibs@stanford.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1718059115/-/DCSupplemental.

is influenced by the level of a second independent variable. Many of these publications have looked for a significant interaction between genetics and the AREDS formulation on AMD progression. Controversy exists based on methodology and subset of patients analyzed. In particular, Awh et al. (12) have stated that genotype groups, defined by combinations of variants in the *CFH* and *ARMS2* genetic regions, can identify individuals who benefit greatly from treatment with the AREDS formulation, as well as those who derive no benefit, or are maybe even harmed. Chew et al. (14) found no evidence of genetic influence on response to AREDS formulation treatment. These analyses included both central GA and NV as clinical end points. Earlier work by Chew et al. (8, 9) noted that AREDS supplements do not delay or prevent central GA, and Seddon et al. (11) found a significant interaction between genetics and AREDS formulation treatment on NV progression but not for central GA. AREDS formulation treatment delays or prevents progression only to NV, and the inclusion of patients who progress to central GA in these analyses dilutes the data and may obscure a significant interaction.

Because the subsequent AREDS2 was designed and conducted without a placebo control arm, no large dataset exists to validate either the primary or any secondary findings from the AREDS. In this study of an expanded dataset from the AREDS, we perform a validation analysis of the interaction of genetics and treatment using bootstrapping, a statistical resampling technique. We used 0.632 bootstrapping to compare the predictive accuracy of models of NV progression risk that may or may not include interaction of genotype group with AREDS formulation treatment. By aggregate analysis of multiple (thousands) random discovery and validation sets generated through resampling of the main dataset, predictors of NV progression can be accurately identified by observing their incremental contribution to model accuracy. This well-established computational method allows powerful statistical determination of the reproducibility of prediction models (15). Validation using the bootstrapping technique can help distinguish false associations, resulting from overfitting or multiple testing, from true ones. Bootstrapping accurately identifies true determinants of clinical outcome better than analysis of a single dataset (15–17).

Materials and Methods

Subjects. Subjects were derived from the AREDS population. Study procedures have been reported previously (8). Subjects were characterized by AREDS investigators at enrollment and during half-yearly follow-ups using retinal images classified by a central reading center. This allowed determination of the time interval from study enrollment to AMD progression to either central GA or NV (8). The AREDS investigators randomized subjects to receive placebo, zinc (80 mg daily), antioxidants (β-carotene, 15 mg; vitamin C, 500 mg; and vitamin E, 400 IU), or both zinc plus antioxidants. Our analyses are restricted to subjects randomized to placebo or to zinc plus antioxidants (the "AREDS formulation"). Subjects who experienced progression events at 2 y or less from study enrollment were not considered in this analysis, since these events were unlikely a result of the assigned treatment (9). The complete phenotype data were provided through the database of Genotypes and Phenotypes (dbGAP) under an investigator agreement with one of the authors (R.K.). This work was approved by the University of Toronto Research Ethics Board. Informed consent was provided by all study subjects upon enrollment in the AREDS.

Genetic Datasets. We assembled genotyping data from three separate sample sources: (*i*) Targeted sequencing was performed by others on 3,340 AMD samples from the Michigan, Mayo, AREDS, Pennsylvania (MMAP) sample set. Short-read sequences were matched to the Genome Reference Consortium build 37 (GRCh37) assembly before being deposited into the NIH's dbGaP database that has been made available through an investigator agreement (R.K.). We obtained the aligned sequences from dbGaP using the NIH's sra-toolkit (version 2.5.4). The read sequences for the *CFH* (chromosome 1) and *ARMS2* (chromosome 10) loci were processed using the Samtools package (www.htslib.org) to deduce unphased genotypes at single-nucleotide polymorphic variants in the complement factor H genomic region (rs1061170, rs3766405, and rs412852) and one SNP (rs10490924) in the age-related maculopathy sensitivity 2 region (www.htslib.org). This yielded genotypes for 2,003 AREDS samples of all presenting grades and treatment groups. (*ii*) Genotyping data were generated from 1,390 AREDS DNA samples purchased from the Coriell Institute (13). Beckman Coulter Genomics according to Good Laboratory Practices performed genotyping at *CFH* and *ARMS2* using bi-directional sequencing. (*iii*) Genotypes at *CFH* rs2755405 and rs412852 and *ARMS2* rs10490924 loci from 534 cases referenced by Chew et al. and collaborators (18, 19) were made available to our group in May 2017 from the National Institutes of Health Office of Research Integrity and Compliance.

Of these samples, we eliminated duplicates and all MMAP samples obtained from subjects who were not part of the AREDS. This resulted in 1,626 samples. Of these, 802 were from subjects randomized to treatment with either placebo or the AREDS formulation, which we refer to as the "expanded" dataset. Of these 802 subjects, 299 had not been part of the prior Awh et al. (12) published analyses. This subgroup of 299 subjects is referred to as the "unique" dataset. Subject distribution between treatment and genetic groups is provided in Fig. 1. AREDS ID/ID2 numbers for each study subject are included as *SI Appendix*, Table 1.

Marker Selection. To analyze the common genetic variability of the *CFH* locus, we selected rs3766405 and rs412852 to tag the two major *CFH* haplotypes as has been done previously (12, 13). We defined the two SNP "high-risk" haplotypes to be rs3766405 CC/rs412852 CC, the average-risk haplotype to be rs3766405 CT/rs412852 CT or rs3766405 CC/rs412852 CT, and the "low-risk" haplotypes to be all other combinations. The derivation of these groups has been described previously (12).

We prespecified genotype groups in the manner described by Awh et al. (12). Briefly, we determined the number of AMD risk alleles at *CFH* and *ARMS2* for each subject. Given the relative rarity of homozygous *CFH* low-risk alleles and *ARMS2* homozygous high-risk alleles, subjects homozygous for these rare alleles were grouped with subjects heterozygous for the corresponding risk alleles (12). Genotype group (GTG) 1 was composed of subjects with low/intermediate *CFH* and no *ARMS2* risk alleles (C01A0). GTG2 subjects had high *CFH* and no *ARMS2* risk alleles (C2A0). GTG3 subjects had low/intermediate *CFH* and one or two *ARMS2* risk alleles (C01A12). GTG 4 subjects had high *CFH* and one or two *ARMS2* risk alleles (C2A12). This is summarized in Table 1.

Clinical Outcome Determination. Subjects in the AREDS cohort varied based on baseline AMD status. Subjects were classified based on the severity of AMD in each eye. We restricted this analysis to subjects with category 3 or 4 AMD at baseline, which are the subgroup of subjects for whom the AREDS formulation was reported beneficial in the original AREDS analysis (9).

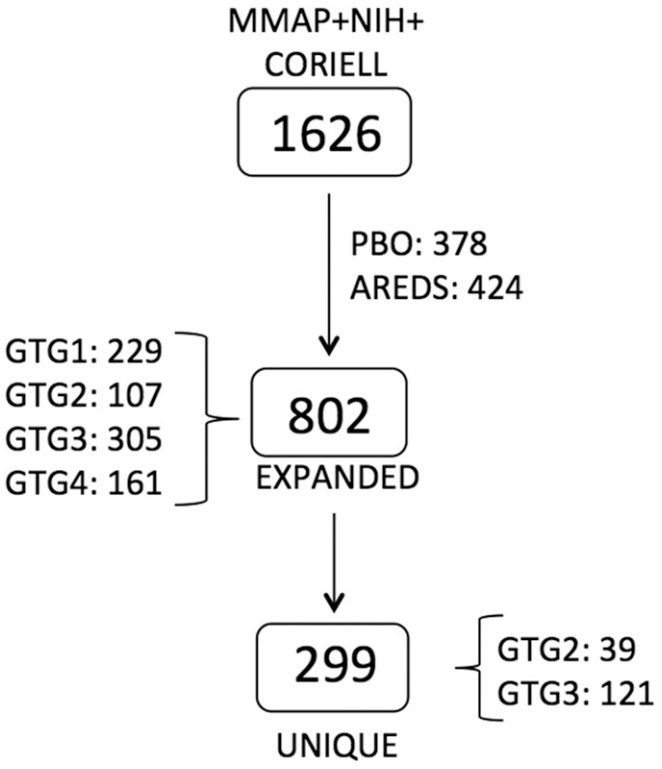

Fig. 1. Distribution of subjects by treatment and genotype group in discovery, validation, and combined sets. PBO, placebo.

Table 1. Previously published genetic grouping using *CFH* and *ARMS2* markers (12)

Genotype group	Subjects (802)	*CFH* risk rs3766405/rs412852	*ARMS2* risk rs10490924
GTG1	229	Low/intermediate-risk all except CC/CC	Low-risk GG
GTG2	107	High-risk CC/CC	Low-risk GG
GTG3	305	Low/intermediate-risk all except CC/CC	High-risk GT/TT
GTG4	161	High-risk CC/CC	High-risk GT/TT

The number of individuals in each group among 802 subjects receiving placebo or AREDS formulation is shown.

Progression of each subject to either NV or central GA was determined through data from within dbGaP tables pht000375.v1.p1.c and pht000376.v1.p1.c1, which provide detailed disease phenotype data for each timed study visit. Definitions for progression to NV or to central GA are documented in "AREDS dbGAP Data Tables: A User's Guide" (20) or in published work from the AREDS retinal image reading center (21). NV was indicated by a score of 11 or 12 in the AMDSEV[R/L]E data field, while central GA by a score of 10 or 12. We have censored observations at 7 y, as has been done in most previous analyses of these data. Since fundus photographs were taken starting at year 2 (20), we also eliminated any progression events reported within the first 2 y, as these were unlikely to be the result of treatment assignment.

Statistical Analysis. The main analyses done in this paper were performed with the expanded dataset, to maximize statistical power. We note that the selection of particular biomarkers and the composition of genotype groupings in Awh et al. (12) were based on a subset of this expanded dataset. This fact could potentially bias our results, causing us to overestimate significance. To guard against this possibility, we have replicated the main results using just the unique set, that is, those subjects who were not used in Awh et al. These validation analyses appear in *Analysis Restricted to the Unique Set*.

Analyses of genetic and nutritional supplement effects and their interactions were done using a Cox proportional hazards model that was adjusted for the following known potential confounders: age at enrollment, sex, and smoking status. Body mass index was not used as a confounder because of a sizable number of missing records. All analysis was done using R statistical software (https://cran.r-project.org) and the rms package (biostat.mc.vanderbilt.edu/wiki/Main/Rrms). Hazard ratios (HRs) and *P* values were calculated using the contrast() function in the rms package to compare specific groups. All statistical code used in this paper is included as *SI Appendix*. This permits the reproduction of all statistical calculations by any investigator with access to AREDS phenotype and genotype datasets for subjects identified in *SI Appendix*, Table 1.

Bootstrap Resampling Validation of Interaction. To evaluate the presence of a genetics–treatment interaction, we used the bootstrapping technique (22, 23). This technique of random sampling with replacement provides a convenient, though computationally intensive, method to make population-wide statistical inferences (Fig. 1). A version called "0.632 bootstrap" is widely used to validate clinical prediction models (24–26). In our setting, the bootstrap exercise was used to assess whether the previously proposed hypothesis of genetics–treatment interaction was a spurious finding or had clinical validity (12). The bootstrap results in this paper were obtained using the R package "pec" (prediction error curves), which uses 0.632 bootstrap for predictive analysis of time-to-event data (27). The statistical models behind the methods used in the pec package utilize the inverse probability of censoring weighted estimators to deal with censored time-to-event data.

We compared event prediction accuracy in subjects using models with and without a genetics–treatment interaction. This approach allows validation of this interaction effect by observing whether prediction models which include the interaction predict the outcome of subjects better than models that do not. Because the purpose of this bootstrap analysis is to test the hypothesis that a genetics–treatment interaction exists, we limited our analysis to the two genotype groups found to be most differently affected by AREDS formulation treatment: individuals GTG2 ($n = 107$) and GTG3 ($n = 305$), for a total of 412 subjects from the expanded set. In each of 100,000 iterations, 412 subjects meeting these criteria were randomly selected with replacement from the 412-subject set. Three Cox models were built using these discovery sets. Study subjects not selected became a paired iteration-specific bootstrap validation sample. On average, 63% of subjects would be selected randomly at least once as a discovery set, for Cox modeling, leaving 37% for bootstrap validation. Three models were considered: (*i*) the base model, which considers genetics, sex, smoking, and age as covariates; (*ii*) the additive model, which considers the presence of AREDS formulation treatment in addition to the base model covariates and assumes that the effect of AREDS formulation treatment is not influenced by genetics; and (*iii*) the interaction model, which considers the same covariates as the additive model but allows for the interaction of AREDS formulation treatment and genetics (Table 2). We also performed the same bootstrap model comparison on all four genotype groups and for subjects found only within the unique set (*SI Appendix*, Table 1).

The predictive accuracy of each model was expressed as the concordance index (C index) (Fig. 2). For a pair of subjects randomly selected from a bootstrap validation set, if one subject experienced progression before the other subject, and the model being evaluated correctly predicted this, that pair of subjects is considered "concordant." The percentage of concordant pairs in each validation set with at least one event and no event-time "ties" is the C index. One C index is generated for each bootstrap iteration and the result is averaged over all 100,000 iterations. This procedure was repeated for a number of time points within the AREDS follow-up time range. For convenience, the C index was converted to a Somers' Dxy measure using the formula Dxy = 2*(C index − 0.5). Both the C index and Somers' Dxy provide the same information as the area under a receiver operator characteristic curve in uncensored data, but Somers' Dxy corrects for random guessing: It is positive if model predictions are better than random guessing, with a maximum value of 1.00 indicating perfect concordance. To generate approximate pointwise 95% confidence intervals, we used quantiles from a block-bootstrap approach (23), where we considered 100,000 bootstrap replications to be 1,000 realizations of bootstrap curves, with each curve based on 100 replications.

Results

Sample Set. Data derived from purchased Coriell AREDS DNA and dbGAP MMAP sequencing and data provided by the NIH Office of Research Integrity and Compliance allowed us to identify 802 AREDS subjects (the expanded dataset) with AREDS category 3 or 4 AMD at study entry treated with either the AREDS formulation or placebo. Of these, we designate 299 subjects not used in the previous Awh analyses as the unique dataset.

Table 2. Covariates in three bootstrap models

Bootstrap model	Description
Base	Considers the effect of genetics and other confounders. This model assumes that progression to NV is not different in subjects treated with AREDS vs. placebo.
Additive	Considers the covariates of the base model, as well as the effect of treatment with the AREDS formulation. Does not allow interaction between genetics and AREDS formulation treatment.
Interaction	Considers the covariates of the additive model, and allows interaction between genetics and AREDS treatment. Assumes that the response to treatment may differ among genotype groups.

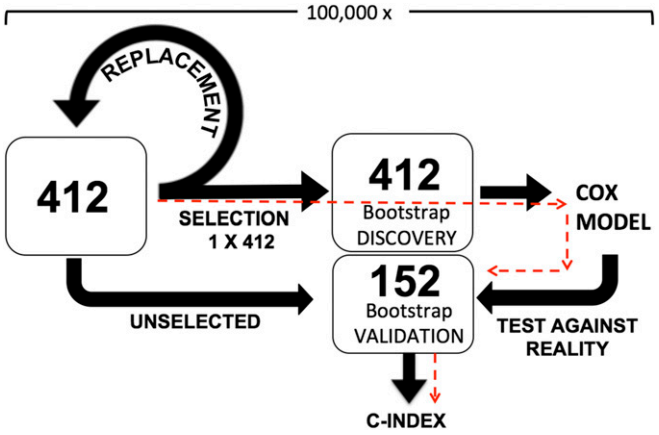

Fig. 2. Bootstrapping schema involving GTG2 and GTG3 subjects from the combined dataset. The number of selections from the expanded set of subjects is the same as the number of available subjects in this series of analyses (412). This selection forms an iteration specific bootstrap discovery set, which contains duplicates because each selected case is returned to the pool before the subsequent case is selected ("replacement"). Subjects not selected at least once for the discovery set are then assigned to the bootstrap validation set (152 on average). A Cox prediction model is generated from the discovery set and tested against the validation set, generating a concordance index. This entire procedure of selection, model derivation from the bootstrap discovery set, and bootstrap set validation is repeated 100 times and the C indices for models built on the same covariates are averaged. To generate 95% confidence intervals, the bootstrap process based on 100 resamplings was repeated 1,000 times. The red arrow shows the order of the process. The C index reflects the predictive power of the Cox covariates. Models including covariates with superior predictive ability will have a higher C index. The addition of uninformative covariates will not result in higher C indices, and will often lower predictive accuracy.

Subjects receiving AREDS formulation treatment and those receiving placebo were balanced with respect to the distribution of *CFH* or *ARMS2* risk alleles, smoking, education level, sex, and age, reflecting random AREDS treatment assignment (data not shown).

***CFH* and *ARMS2* and AREDS Formulation Association with NV and Central GA.** We first performed an additive Cox regression to evaluate the effect of *CFH* and *ARMS2* alleles on the risk of progression to neovascular AMD and on the risk for developing central geographic atrophy within the expanded set of 802 subjects. The adjusted HR for NV and central GA for subjects having two high-risk *CFH* alleles compared with intermediate- and low-risk subjects was 1.72 ($P = 0.0024$) and 1.26 ($P = 0.26$), respectively. Those with high-risk *ARMS2* genotypes had an HR of 2.76 ($P < 0.0001$) for developing NV and an HR of 1.65 ($P = 0.03$) for developing GA.

Interaction of Genetics and Treatment Group with Progression to NA or GA. We used the Cox regression function in data from 802 subjects in the expanded dataset to examine statistical interactions between treatment group (AREDS formulation or placebo) and *CFH* and *ARMS2* genotype group on the progression to NV and central GA separately. Strong interaction was seen between *CFH* and *ARMS2* risk alleles and AREDS formulation treatment with NV as a progression end point (Table 3). No significant interaction was observed between AREDS formulation treatment and *CFH* or *ARMS2* risk alleles on progression to central GA. Given these observations, we confined subsequent analyses of genetics–treatment interaction to its effect on progression to NV.

To test the clinical utility of these observations for making individual treatment recommendations, we generated Cox model adjusted survival curves which control for age, sex, and smoking as covariates. For subjects with high *CFH* risk alleles and without *ARMS2* risk alleles (GTG2), a significant increase in the rate of progression to NV is observed among those treated with the AREDS formulation compared with placebo (HR = 2.9, $P = 0.018$) (Fig. 3, *Left*). For subjects with the opposite genetic risk pattern, low *CFH* and high *ARMS2* risk alleles (GTG3), a marked decrease in the rate of progression to NV was observed among those treated with the AREDS formulation vs. placebo (HR = 0.50, $P = 0.008$) (Fig. 3, *Right*).

Bootstrapping Validation. To further test for the presence of a genetics–treatment interaction effect in the AREDS population, we evaluated the relative predictive accuracy of models of progression risk for NV that either include an interaction effect or do not. The pointwise 95% confidence bands strongly suggest the superior predictive ability of the interaction model over the other two models across later time points (4.5 y and beyond). Equal predictive performance was observed for the base model (genetics with demographics) and the additive model, which adds knowledge of AREDS formulation treatment but without allowing genetic interaction. This suggests a negligible average treatment effect of the AREDS formulation when interaction with genetics is not considered (Fig. 4). Similar bootstrap results generated both from the full expanded set of 802 cases of all four genotype groups and from subjects in the unique set are provided (*SI Appendix*, Figs. 1 and 2).

Analysis Restricted to the Unique Set. A concern of data overfitting has been raised in response to publications by Awh et al. (12, 18, 19) that first described the relationship between nutritional treatment and *CFH*/*ARMS2* genotype combinations. Selection of particular biomarkers and the composition of genetic groupings may have resulted in inflated statistical significance. As a validation analysis, we replicated main genetics–treatment interaction findings in subjects from the unique set, consisting of subjects who were not part of any previous analysis by Awh et al. (12, 18).

A Cox regression analysis of the unique dataset (placebo or AREDS formulation-treated; $n = 299$) was adjusted for age, sex, and smoking, in the same fashion as in the expanded set. Due to the comparatively small sample size, time censoring was not performed to maximize the number of NV progression events available for analysis. We observe that subjects with high *CFH* risk alleles and no *ARMS2* risk alleles (GTG2; $n = 82$) have a significant increase in progression risk if treated with the AREDS formulation vs. placebo (Table 4). The opposite response was observed among subjects with low *CFH* risk alleles and high *ARMS2* risk alleles (GTG3; $n = 238$) (Fig. 5 and *SI Appendix*, Table 2). These subjects had a significant reduction in AMD progression risk if treated with the AREDS formulation.

Table 3. Hazard ratios of AREDS formulation treatment for progression to NV or central GA in four genotype groups (with *P* values), and *P* values of tests for interaction of genotype group and AREDS formulation treatment effect for the expanded dataset ($n = 802$)

	AREDS-NV		AREDS-GA	
Genotype group	HR	*P* value	HR	*P* value
GTG1	1.41	0.43	0.70	0.40
GTG2	2.92	0.018	1.04	0.93
GTG3	0.50	0.008	0.60	0.09
GTG4	1.03	0.91	0.88	0.74
Interaction (χ^2, 2 df)	—	0.01	—	0.62

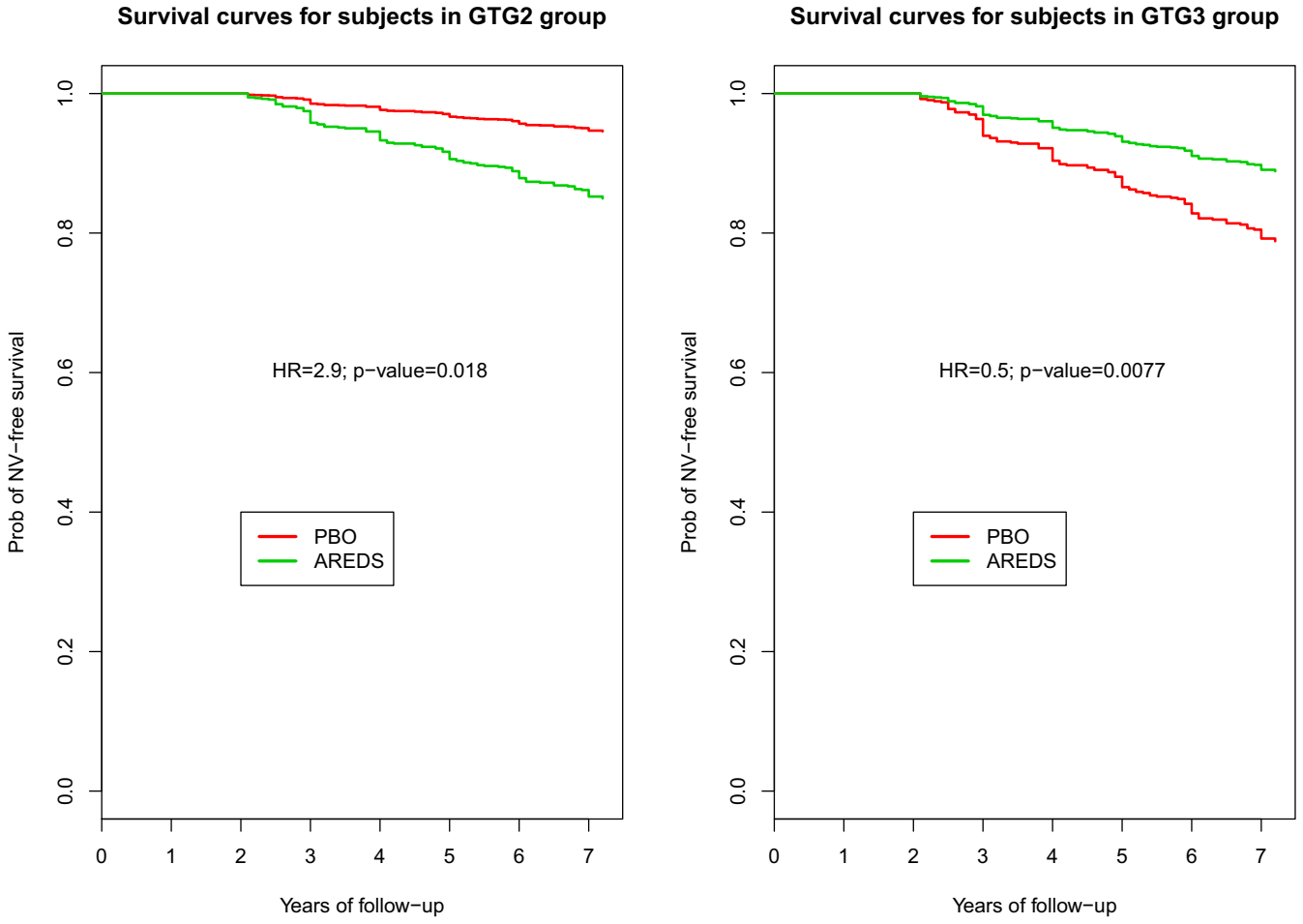

Fig. 3. Cox-derived survival curves using expanded datasets for NV-free survival for subjects with high *CFH* and no *ARMS2* risk alleles (GTG2; $n = 107$) (*Left*) and for individuals with low *CFH* and high *ARMS2* risk (GTG3; $n = 305$) (*Right*). Subjects in both panels were treated with either placebo or the AREDS formulation.

Discussion

Our analyses of the most comprehensive collection of AREDS data and DNA to date show that the risk of progression from intermediate to neovascular AMD is significantly altered by AREDS formulation treatment in a genotype group-dependent fashion, confirming prior independent reports of a significant interaction between AREDS formulation treatment, genetic risk, and progression to advanced AMD (11, 12).

Klein et al. (10) first observed that the benefit of the AREDS formulation was eliminated for subjects with high-risk *CFH* alleles. The authors postulated that this effect was due to the high-dose zinc component of the AREDS formulation. Awh et al. (12) analyzed the response to AREDS nutritional supplements as influenced by *CFH* and *ARMS2* genetic risk in 989 AREDS subjects and confirmed the observation of Klein et al. with regard to *CFH*, while identifying an opposite interaction with *ARMS2* polymorphisms. Subjects with high *CFH* and low *ARMS2* risk alleles had increased AMD progression if treated with zinc (alone, or as a component of the AREDS formulation), while those with low *CFH* and high *ARMS2* risk alleles had decreased progression (13). Chew et al. (14) published a statistical analysis of 1,237 AREDS subjects and found no influence of genetics on response to the AREDS formulation. However, their analyses were performed separately on 27 relatively small genetic risk–treatment groups, a design statistically underpowered to demonstrate any interaction. Seddon et al. (11) analyzed progression to overall advanced AMD and progression to the two subtypes of advanced AMD, NV and central GA. They found that for subjects with low *CFH* and high *ARMS2* genetic risk, the reduction in overall advanced AMD was due to decreased progression to NV, with no significant effect on central GA progression. These authors concluded that "the effectiveness of antioxidant and zinc supplementation appears to differ by genotype" (11). Their approach differs from ours in that they considered all subjects regardless of baseline AMD status, used a single eye as a unit of observation, considered a single SNP to tag a *CFH* region rather than the two-SNP-based haplotype we use, and did not analyze the GTG2 group separately. Despite these differences, they also concluded that the genetics–treatment interaction predicts progression to NV and not to central GA.

We assembled a dataset derived from the MMAP archive, Coriell samples, and NIH Office of Intramural Research Integrity and Compliance. We performed two validation studies to address the potential of overfitting and spurious findings in previous studies: (*i*) a bootstrap predictive validation of genetics–treatment interaction model; and (*ii*) validation of the results in a unique validation subsample that does not include subjects used previously in defining genotype groups. Our analysis of this expanded dataset supports prior observations that the effect of AREDS formulation treatment on progression to advanced AMD is driven by changes in the risk of developing NV, not by changes in the risk of developing central GA. We confirm that individuals with high *CFH* and no *ARMS2* risk alleles have an increased risk of progression to NV if treated with the AREDS formulation compared

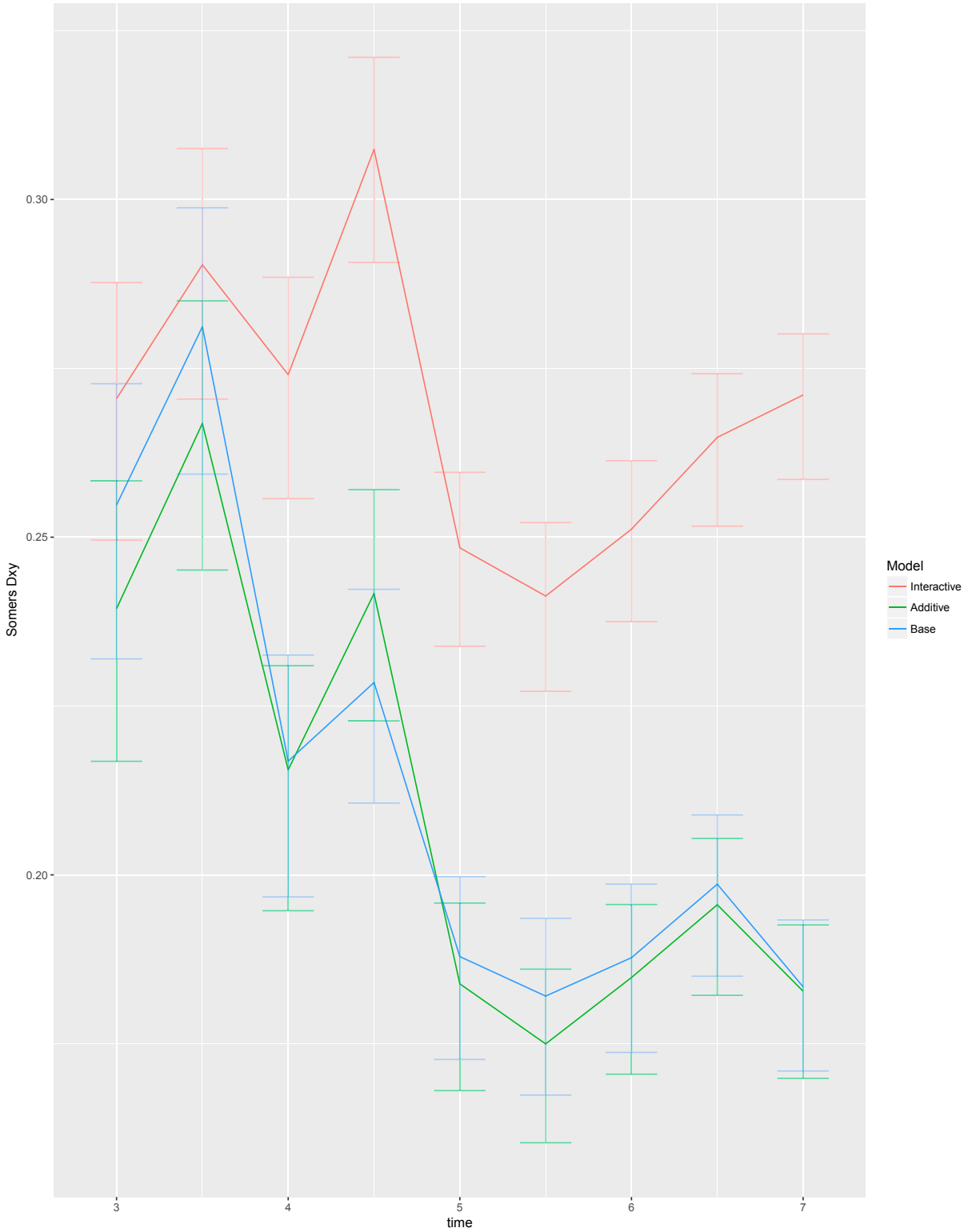

Fig. 4. A 0.632 bootstrap analysis of GTG2 and GTG3 subjects from the expanded dataset (n = 412) showing superior prediction of NV progression for the interaction model. Somers' Dxy (a calibrated C-index measure) for three different Cox proportional hazards models was generated across follow-up time points (x) with approximate pointwise 95% confidence intervals. The interaction model includes sex, smoking, age, AREDS formulation, genetics, and interaction between AREDS formulation treatment and genetics. The additive model includes all of the covariates in the interaction model, but does not allow for interaction. The base model considers only sex, smoking, age, and genetics as covariates.

with placebo. For individuals with low *CFH* and high *ARMS2* risk alleles, we found a substantial beneficial effect from AREDS formulation treatment. These findings are consistent with those of Awh et al. (12) and Seddon et al. (11), even though our analysis differed from theirs in the following ways: (*i*) We considered progression to NV as the relevant event rather than progression to

Table 4. Hazard ratios and *P* values (Wald) for genotype group and treatment with the AREDS formulation vs. placebo on the development of NV for the 299 subjects in the unique set (12)

Genotype group	HR AREDS vs. placebo	*P* value
GTG2	4.9	0.021
GTG3	0.36	0.003

either NV or central GA; and (*ii*) we restricted analysis to subjects treated with either the AREDS formulation or placebo. The decision to analyze progression to only NV is supported by published evidence that the AREDS formulation is effective only in the prevention of the NV form of advanced AMD (original AREDS reports 8 and 35) (8, 9) and by our analysis of the interaction of GTG and progression to NV or central GA (Table 2). Insensitivity to this clinical distinction contributed to an inaccurate conclusion by Assel et al. (27) that the AREDS formulation is beneficial in genetically unselected individuals. This contributed to their inability to show that individuals in GTG2 treated with the AREDS supplementation have an increased risk of progression to NV. We also decided to analyze only those subjects treated with either the AREDS formulation or placebo, excluding those treated with only antioxidants or only zinc, to focus our analysis on the "real-world" decision confronting most patients and physicians: to treat or to not treat.

Our validation analysis of a unique subgroup of AREDS formulation- or placebo-treated subjects also confirms the genetics–treatment interaction found by Seddon et al. and Awh et al. (11, 12). This is a group of subjects whose data had not been used to derive the prespecified genotype groups, hence eliminating the potential for data overfitting and spurious findings (12). Our analysis differs from a report by Chew et al., in which the authors were unable to validate the findings of Awh et al. (12, 18). The number of relevant cases in our validation dataset is larger because of the additional cases available through MMAP. We also evaluated progression only to NV and not to central GA, since progression to central GA has not been shown to be affected by supplementation in the original AREDS study.

As a second form of validation, we performed a bootstrap analysis of genetics–treatment interaction by comparing the predictive ability of a model that considers genetics–treatment interaction with models that do not. This validation is distinct and more stringent than the conclusions derived from a Cox model fitted to a single sample. Spurious covariates identified in a derivation dataset would not improve the prediction of the outcome of subjects external to the derivation dataset, regardless of their initial apparent statistical significance. In this sense, bootstrap validation could be used more broadly to help identify important interactions, although here we use it strictly to help confirm the significance of a previously suggested genetics–treatment interaction. Our bootstrap predictive validation illustrates the clinical significance of treatment and genetics interaction in predicting

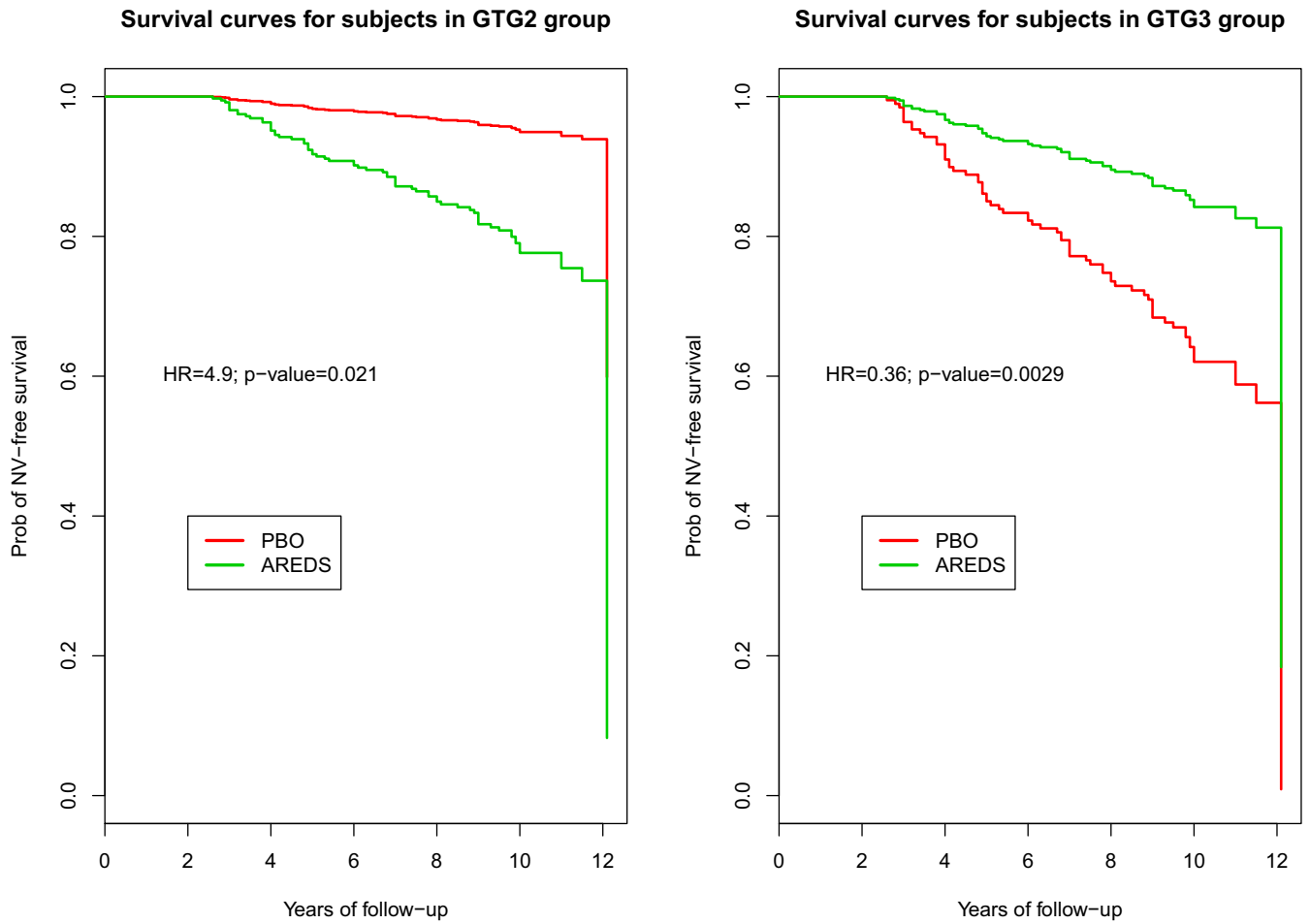

Fig. 5. Cox model-derived survival curves of NV-free survival for unique set subjects with high *CFH* risk alleles and no *ARMS2* risk alleles (GTG2; $n = 39$) (*Left*) and low *CFH* risk alleles and high *ARMS2* risk alleles (GTG3; $n = 121$) (*Right*). Subjects were treated with placebo or the AREDS formulation.

the risk of progression events in new subjects. We considered three competing models that each attempts to explain the risk of progression to NV. The base model includes genetics, smoking, age, and sex as predictors of AMD progression. The additive model adds the effect of AREDS formulation treatment but does not allow any effect modification by genetics on AREDS treatment (i.e., excludes the possibility of interaction). The interaction model builds on the additive model by allowing a genetics–AREDS formulation treatment interaction. Remarkably, the additive model shows only a negligible improvement in prediction ability over the base model, despite the addition of information regarding AREDS formulation treatment status. When the genetics–AREDS formulation treatment interaction effect is allowed, the interaction model shows significant improvement in accuracy (Fig. 4). We see that the predictive accuracy for all three models is similar in earlier years, diverging at and beyond 4 y post randomization. This suggests that any treatment effect accrues over time. The superior predictive performance of the interaction model confirms that the effectiveness of AREDS formulation treatment is dependent upon *CFH* and *ARMS2* genetic risk status. The conclusions from this primary bootstrap analysis on all subjects also hold when the analysis is performed only within the unique set [i.e., subjects not analyzed by Awh et al. (12)], further demonstrating that the previously defined genotype groups and their interactions with treatment were not spurious effects.

The clinical utility of this observation is illustrated when progression is considered as a function of genetics using Cox survival estimates (Fig. 3), which shows that one genotype group (GTG2) is likely harmed by treatment with the AREDS formulation, while the other (GTG3) is likely to substantially benefit from treatment. It appears that the overall modest benefit of the AREDS formulation is the result of effects in a genotype group that does extraordinarily well, a genotype group that does worse than with placebo, and subjects whose outcome is relatively unaffected by treatment with the AREDS formulation (GTG1 and GTG4).

While the AREDS study defined AMD as a posterior pole disease, it may not be strictly limited to the posterior pole, as drusen and peripheral pigmentary changes are common in the periphery of AMD eyes (28, 29). Genetic associations with peripheral phenotypes (30, 31) further suggest that AMD may not be limited to posterior pathology and, if confirmed, could also be valuable in helping determine the appropriate prophylactic nutritional formulation for an expanded number of AMD patients.

The human retina concentrates zinc, with levels influenced by chronic oral zinc supplementation (32). Biochemical analysis of the interaction of complement proteins and zinc shows concentration-dependent oligomerization or insolubility, suggesting a potential defense against uncontrolled activation (33–36). Pathologic variations in complement components may interact with zinc. For instance, an AMD-associated polymorphism in complement component 3 impedes binding to its inhibitor, complement factor H, preventing the formation of a complex highly sensitive to zinc-induced inactivation (37). AMD-associated *CFH* allotypes are less able to inactivate *C3b* due to impeded binding to the acute-phase reactant C-reactive protein, which bridges these two inflammatory regulators (34). Through such mechanisms, polymorphisms in *CFH* may alter the normal role of zinc inactivation of the complement cascade and may provide the physiologic basis for our observations.

Our observations regarding genetic risk, AMD, and treatment with the AREDS formulation are based upon multiple statistical analyses of DNA and data from one of the largest groups of AREDS subjects yet assembled. They confirm the existence of a genetics–treatment interaction identified by multiple independent researchers (10–12). They validate that the response to AREDS formulation treatment varies substantially among individuals, based on *CFH* and *ARMS2* genetic risk. They show that the genotype groups previously reported may be an effective method of identifying individuals who are likely to benefit, or not, from treatment (12). There is no placebo-controlled replication study of the AREDS to be used for additional validation. However, the lack of a replication trial has not prevented the widespread acceptance and recommendation of the AREDS formulation treatment for subjects with intermediate AMD. Bootstrap validation demonstrates that the genetics–treatment interaction is consistent and valid within this dataset, providing support for the clinical validity of the interaction of *CFH* and *ARMS2* genotype groups with AREDS formulation treatment.

ACKNOWLEDGMENTS. The authors received no financial support for the research, authorship, or publication of this article.

1. Pennington KL, DeAngelis MM (2016) Epidemiology of age-related macular degeneration (AMD): Associations with cardiovascular disease phenotypes and lipid factors. *Eye Vis (Lond)* 3:34.
2. Schmier JK, Covert DW, Lau EC (2012) Patterns and costs associated with progression of age-related macular degeneration. *Am J Ophthalmol* 154:675–681.e1.
3. Jager RD, Mieler WF, Miller JW (2008) Age-related macular degeneration. *N Engl J Med* 358:2606–2617.
4. Seddon JM, Reynolds R, Yu Y, Daly MJ, Rosner B (2011) Risk models for progression to advanced age-related macular degeneration using demographic, environmental, genetic, and ocular factors. *Ophthalmology* 118:2203–2211.
5. Lechanteur YT, et al. (2015) Association of smoking and CFH and ARMS2 risk variants with younger age at onset of neovascular age-related macular degeneration. *JAMA Ophthalmol* 133:533–541.
6. Seddon JM, et al. (2007) Association of CFH Y402H and LOC387715 A69S with progression of age-related macular degeneration. *JAMA* 297:1793–1800.
7. Yu Y, Reynolds R, Rosner B, Daly MJ, Seddon JM (2012) Prospective assessment of genetic effects on progression to different stages of age-related macular degeneration using multistate Markov models. *Invest Ophthalmol Vis Sci* 53:1548–1556.
8. Age-Related Eye Disease Study Research Group (2001) A randomized, placebo-controlled, clinical trial of high-dose supplementation with vitamins C and E, beta carotene, and zinc for age-related macular degeneration and vision loss: AREDS report no. 8. *Arch Ophthalmol* 119:1417–1436.
9. Chew EY, et al. (2013) Long-term effects of vitamins C and E, beta-carotene, and zinc on age-related macular degeneration: AREDS report no. 35. *Ophthalmology* 120: 1604–1611.e4.
10. Klein ML, et al. (2008) CFH and LOC387715/ARMS2 genotypes and treatment with antioxidants and zinc for age-related macular degeneration. *Ophthalmology* 115: 1019–1025.
11. Seddon JM, Silver RE, Rosner B (2016) Response to AREDS supplements according to genetic factors: Survival analysis approach using the eye as the unit of analysis. *Br J Ophthalmol* 100:1731–1737.
12. Awh CC, Hawken S, Zanke BW (2015) Treatment response to antioxidants and zinc based on CFH and ARMS2 genetic risk allele number in the Age-Related Eye Disease Study. *Ophthalmology* 122:162–169.
13. Awh CC, Lane AM, Hawken S, Zanke B, Kim IK (2013) CFH and ARMS2 genetic polymorphisms predict response to antioxidants and zinc in patients with age-related macular degeneration. *Ophthalmology* 120:2317–2323.
14. Chew EY, et al. (2014) No clinically significant association between CFH and ARMS2 genotypes and response to nutritional supplements: AREDS report number 38. *Ophthalmology* 121:2173–2180.
15. Efron B (1979) Bootstrap methods: Another look at the jackknife. *Ann Stat* 7:1–26.
16. Brunelli A, Rocco G (2006) Internal validation of risk models in lung resection surgery: Bootstrap versus training-and-test sampling. *J Thorac Cardiovasc Surg* 131:1243–1247.
17. Kulesa A, Krzywinski M, Blainey P, Altman N (2015) Sampling distributions and the bootstrap. *Nat Methods* 12:477–478.
18. Chew EY, Klein ML, Clemons TE, Agrón E, Abecasis GR (2015) Genetic testing in persons with age-related macular degeneration and the use of the AREDS supplements: To test or not to test? *Ophthalmology* 122:212–215.
19. Wittes J, Musch DC (2015) Should we test for genotype in deciding on Age-Related Eye Disease Study supplementation? *Ophthalmology* 122:3–5.
20. Henning A (2009) AREDS dbGAP Data Tables: A User's Guide. Available at https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/GetPdf.cgi?id=phd001552.1. Accessed November 13, 2017.
21. Davis MD, et al.; Age-Related Eye Disease Study Group (2005) The age-related eye disease study severity scale for age-related macular degeneration: AREDS report no. 17. *Arch Ophthalmol* 123:1484–1498.
22. Efron B, Tibshirani R (1991) Statistical data analysis in the computer age. *Science* 253: 390–395.
23. Efron B, Tibshirani RJ (1993) *An Introduction to the Bootstrap* (Chapman and Hall/CRC, New York).
24. Ross JS, et al. (2010) Hospital volume and 30-day mortality for three common medical conditions. *N Engl J Med* 362:1110–1118.

25. Henderson AR (2005) The bootstrap: A technique for data-driven statistics. Using computer-intensive analyses to explore experimental data. *Clin Chim Acta* 359:1–26.
26. Gerds TA, Kattan MW, Schumacher M, Yu C (2013) Estimating a time-dependent concordance index for survival prediction models with covariate dependent censoring. *Stat Med* 32:2173–2184.
27. Assel MJ, et al. (October 9, 2017) Genetic polymorphisms of CFH and ARMS2 do not predict response to antioxidants and zinc in patients with age-related macular degeneration: Independent statistical evaluations of data from the Age-Related Eye Disease Study. *Ophthalmology*, 10.1016/j.ophtha.2017.09.008.
28. Domalpally A, et al.; Writing Committee for the OPTOS PEripheral RetinA (OPERA) study (Ancillary Study of Age-Related Eye Disease Study 2) (2017) Peripheral retinal changes associated with age-related macular degeneration in the Age-Related Eye Disease Study 2: Age-Related Eye Disease Study 2 report number 12 by the Age-Related Eye Disease Study 2 Optos PEripheral RetinA (OPERA) study research group. *Ophthalmology* 124:479–487.
29. Lengyel I, et al. (2015) A population-based ultra-widefield digital image grading study for age-related macular degeneration-like lesions at the peripheral retina. *Ophthalmology* 122:1340–1347.
30. Shuler RK, Jr, et al. (2008) Peripheral reticular pigmentary change is associated with complement factor H polymorphism (Y402H) in age-related macular degeneration. *Ophthalmology* 115:520–524.
31. Seddon JM, Reynolds R, Rosner B (2009) Peripheral retinal drusen and reticular pigment: Association with *CFHY402H* and *CFHrs1410996* genotypes in family and twin studies. *Invest Ophthalmol Vis Sci* 50:586–591.
32. Lengyel I, et al. (2007) High concentration of zinc in sub-retinal pigment epithelial deposits. *Exp Eye Res* 84:772–780.
33. Perkins SJ, Nan R, Li K, Khan S, Miller A (2012) Complement factor H-ligand interactions: Self-association, multivalency and dissociation constants. *Immunobiology* 217: 281–297.
34. Perkins SJ, Okemefuna AI, Nan R (2010) Unravelling protein-protein interactions between complement factor H and C-reactive protein using a multidisciplinary strategy. *Biochem Soc Trans* 38:894–900.
35. Nan R, et al. (2011) Zinc binding to the Tyr402 and His402 allotypes of complement factor H: Possible implications for age-related macular degeneration. *J Mol Biol* 408: 714–735.
36. Nan R, Gor J, Lengyel I, Perkins SJ (2008) Uncontrolled zinc- and copper-induced oligomerisation of the human complement regulator factor H and its possible implications for function and disease. *J Mol Biol* 384:1341–1352.
37. Rodriguez E, Nan R, Li K, Gor J, Perkins SJ (2015) A revised mechanism for the activation of complement C3 to C3b: A molecular explanation of a disease-associated polymorphism. *J Biol Chem* 290:2334–2350.