AMERICAN ACADEMY OF OPHTHALMOLOGY®

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

Melissa J. Assel, MS,[1,*] *Fan Li, MS,*[2,*] *Ying Wang, PhD,*[3,*] *Andrew S. Allen, PhD,*[2,‡] *Keith A. Baggerly, PhD,*[3,‡] *Andrew J. Vickers, PhD*[1,‡]

Purpose: Considerable controversy has erupted in recent years regarding whether genotyping should be part of standard care for patients with age-related macular degeneration (AMD) who are being considered for treatment with antioxidants and zinc. We aimed to determine whether genotype predicts response to supplements in AMD.

Design: Three separate statistical teams reanalyzed data derived from the Age-Related Eye Disease Study (AREDS), receiving data prepared by the AREDS investigators and, separately, data from investigators reporting findings that support the use of genotyping.

Participants: The population of interest was AREDS participants with AMD worse than category 1 and genotyping data available. Data from the 2 groups overlap imperfectly with respect to measurements made: the largest common set involved 879 participants for whom the same *CFH* and *ARMS2* single nucleotide polymorphisms were measured by both groups.

Methods: Each team took a separate but complementary approach. One team focused on data concordance between conflicting studies. A second team focused on replicating the key claim of an interaction between genotype and treatment. The third team took a blank slate approach in attempting to find baseline predictors of treatment response.

Main Outcome Measures: Progression to advanced AMD.

Results: We found errors in the data used to support the initial claim of genotype–treatment interaction. Although we found evidence that high-risk patients had more to gain from treatment, we were unable to replicate any genotype–treatment interactions after adjusting for multiple testing. We tested 1 genotype claim on an independent set of data, with negative results. Even if we assumed that interactions in fact did exist, we did not find evidence to support the claim that supplementation leads to a large increase in the risk of advanced AMD in some genotype subgroups.

Conclusions: Patients who meet criteria for supplements to prevent AMD progression should be offered zinc and antioxidants without consideration of genotype. *Ophthalmology 2018;125:391-397 © 2017 by the American Academy of Ophthalmology*

Supplemental material available at www.aaojournal.org.

The Age-Related Eye Disease Study (AREDS) was a large, multicenter, double-blind randomized trial to determine whether high-dose antioxidants, zinc, or their combination could reduce the risk of progression of age-related macular degeneration (AMD) in older patients. Excluding patients in AMD category 1, for whom the event rate was less than 1%, the combination of zinc and antioxidants was found to reduce the risk of progression to advanced AMD (odds ratio, 0.68; 95% confidence interval [CI], 0.49–0.93; $P = 0.002$).[1] The publication of the trial results led to rapid changes in practice, with at-risk patients routinely prescribed the zinc and antioxidant combination tested in the trial.

In 2008, Klein et al[2] published a pharmacogenomic study suggesting that the effects of antioxidants and zinc on AMD in AREDS may be influenced by genotype,

https://doi.org/10.1016/j.ophtha.2017.09.008
ISSN 0161-6420/17

specifically, the disease-related genes age-related maculopathy susceptibility 2 (*ARMS2*) and complement factor H (*CFH*), also known as *ARMS1*. For instance, there was a smaller difference between treatment and placebo in patients with the CC genotype for *CFH* Y402H (44% vs. 39%) compared with those with the TT genotype (34% vs. 11%; $P = 0.03$ for interaction). No interaction was found for *LOC387715/ARMS2*. The authors made only cautious conclusions, stating that "corroboration … is needed before considering modification of current management." Such corroboration seemed to come from Awh et al,[3] who examined the relative benefit of treatment across a wider set of genotypes from 11 disease-related markers before settling on 2 markers for *CFH* and 1 marker for *ARMS2*. Importantly, Awh et al claimed qualitative interactions between genotype and treatment outcome. The authors stated that the "data support a deleterious interaction between CFH risk alleles and high-dose zinc supplementation" such that patients with certain genotypes should be treated by antioxidants alone rather than by antioxidants plus zinc. The conclusions included "recommendations" that would lead to "improved outcomes through genotype-directed therapy."

These findings led the original study authors, Chew et al,[4] to attempt a replication. Measuring the genotype of a different subset of patients from AREDS, the authors did find the anticipated prognostic relationship between *CFH* and *ARMS2* genotype and risk of progression. However, they did not find any predictive relationship between genotype and treatment effect, with test results for interaction being nonsignificant. The authors concluded that "supplements reduced the rate of AMD progression across all genotype groups" and that genetic testing should not be used to determine treatment. These negative findings were challenged by Awh and Zanke,[5] who claimed that the study by Chew et al refutes any claim of overall benefit for supplementation and that a separate editorial, written by a well-known statistician and epidemiologist team (Wittes and Musch[6]) supported the genotyping. In response, Chew et al[7] claimed that Awh and Zanke had misinterpreted their study and that, in fact, the Wittes and Musch editorial favored their own position.

To help resolve this debate, the Office of Intramural Research at the National Institutes of Health (NIH) asked our 3 biostatistical groups to re-examine independently the data used by Awh et al[3] and Chew et al[5] to determine whether genotyping should be part of the clinical decision of whether to use supplements for AMD prevention. Herein, we report our findings.

Methods

A research integrity officer at the NIH contacted both sets of investigators (Chew et al and Awh et al) and proposed that they provide data to be forwarded on to independent biostatisticians—whose names and affiliations were not revealed—for further analysis. The 2 groups agreed and sent their data to the research integrity officer, who forwarded it on to us. Neither the NIH nor any other outside group or investigator participated in the design of the statistical methods used, interpretation of the results, drafting of the manuscript, or manuscript review before submission. No direct funding or any other type of financial remuneration was provided by NIH to support the current work.

Clinical information on AREDS participants is available to qualified researchers through the Database of Genotypes and Phenotypes, and single nucleotide polymorphism (SNP) and sequencing data are available now for an ever-increasing subset, although significantly fewer data were available when the debate began. For their studies, Awh et al[3,8] focused on 979 patients for whom blood samples could be obtained from the Coriell biorepository. They used these samples to perform their own genotyping. They genotyped *CFH* at 2 SNPs, rs3766405 and rs412852, and assessed insertion/deletion (indel) status for *ARMS2* at 1 location. Chew et al[4,9] looked at data from 1237 patients for whom they had *CFH* and *ARMS2* genotype data at SNPs other than those used by Awh et al (rs1061170 and rs1410996 for *CFH* and rs10490924 for *ARMS2*; summarized in their Fig 1B) and from 1413 patients measured using exactly the same SNPs as those used by Awh et al (summarized in their Fig 1C). In all, genotype data from these 3 locations from Awh et al are available for 1523 participants: 879 were measured by both groups, 110 were measured only by Awh et al, and 534 were measured only by Chew et al. All data can be matched using anonymized AREDS patient identifiers.

The genotype data for patients measured at the above mentioned 3 SNPs underwent several levels of summarization. First, there were the raw genotype assessments (AA, AB, or BB) at each of the 3 SNPs. Second, results were expressed at the gene level in terms of the number of risk alleles for that gene (0, 1, or 2). This mapping is straightforward for *ARMS2* (measured at just 1 SNP), but requires more detailed specification for *CFH* to indicate how a pair of genotypes is reduced to a number. Third, the numbers of risk alleles for each of the 2 genes are used to assign patients to genotype groups (GTGs). Proposed treatment differentiation would occur at the GTG level.

The 3 statistical groups decided to work independently on 3 separate approaches to the replication problem. The MD Anderson Cancer Center group focused primarily on checking data and evaluating concordance between different data sets. The Duke University group's role was to replicate the key findings of Awh et al concerning interactions between genotype and outcome. The Memorial Sloan Kettering Cancer Center (MSKCC) group took a blank slate approach, using all baseline data, including both clinical variables and genotype data, to determine whether benefit from treatment could be predicted.

MD Anderson Cancer Center: Data Concordance

We received raw data on patients from AREDS[1] linking times to AMD disease progression to *CFH* and *ARMS2* genotypes and treatment group, from both Awh et al (Arctic)[3,8] and the AREDS investigators.[4,9] The data also contained various clinical covariates such as age, gender, race, body mass index (BMI), and smoking history.

Because unappreciated differences between data sets could explain some of the published inconsistencies, first we extensively checked the raw data supplied by both groups. We cross-tabulated genotype calls for rs3766405, genotype calls for rs412852, and the reported numbers of *CFH* risk alleles. We also checked progression data in each of the 2 data sets by examining the longitudinal data on AMD eye categories to identify the time point at which either progression to category 4 in either eye first occurs, if the patient's category values were less than 4 for both eyes at the outset, or progression to category 4 occurs in the non–category 4 eye if the patient has 1 eye rated as category 4 at the outset.

We split the data into 3 groups based on whether we had genotype call data for *CFH* at rs3766405 and rs412852 and indel data for *ARMS2* from both Arctic and AREDS (879 patients), just Arctic (110 patients), or just AREDS (534 patients). Each of the 3 data sets contains AREDS identification, *CFH* genotype, *ARMS2* genotype, treatment group index, and progression status and time from each of the 2 groups. The cleaned data were presented to and approved by all 3 groups before they started their own independent analyses.

As noted by Chew et al[4] in 2014, the longer follow-up times now available include times after the end of randomization for the initial trial, at which point the all different treatment groups were shifted to receive the AREDS formulation. Because this could distort treatment differences, we chose to work with the progression-free survival (PFS) data from AREDS, which were censored at the end of 2001. Using the raw genotype calls at all SNPs, we assigned patients to gene severity levels and GTGs. Then, using GTG information, we used Cox proportional hazards models to fit time to disease progression as a function of various covariates, using the samples measured by both groups, and checked whether terms identified as significant retained their importance in the data sets examined by just 1 group. For this, we focused on GTG 2 (both *CFH* SNPs are CC, and *ARMS2* category is 11) from Awh et al[8] in 2015, because this was the subgroup for which the strongest claims were made. We started with the 120 GTG 2 patients examined by both Arctic and AREDS investigators. Then, as a validation test, we performed the same analyses using the 75 GTG 2 patients examined by the AREDS group alone. This is similar to the approach taken by Chew et al[9] in 2015 but uses the direct matching we were able to obtain with access to data sets from both groups that Chew et al did not have. Analyses were conducted using R software version 2.3.0 (R Foundation for Statistical Computing, Vienna, Austria).

Duke University: Interaction between Genotype and Treatment

We included in the analysis 879 AREDS patients who have a high risk of progression to late AMD and who are included in both the Awh et al and Chew et al analyses. Among the 879 AREDS patients, 673 patients had intermediate AMD in 1 or both eyes (AREDS AMD category 3) and 206 patients had late AMD in 1 eye (AREDS AMD category 4).

We determined whether there was an interaction between the genotypes and treatment with antioxidant plus zinc for PFS. All patients were followed up in the randomized controlled trial until the end of 2001, and additional follow-up data are available through 2005 in an observational study. Because potential treatment noncompliance or crossover effects in the observational phase of study might have introduced bias, we confined our analysis to the period of the randomized controlled trial in 2001.

A comprehensive set of 11 genetic markers for AMD was identified by Awh et al[3] (Table 1) that we used for our primary analysis. Following Awh et al, homozygous minor allele counts were combined with heterozygotes for markers with low minor allele frequencies (<1%). Because the process of summarizing *CFH* genotypes (rs3766405, rs412852) into risk allele counts was applied more consistently in the data from Chew et al[4] than in the data from Awh et al,[3] we used the risk allele counts from Chew et al.[7] The *ARMS2* risk allele counts (372_815de1443ins54) were highly concordant between the 2 studies, and we used the counts from Chew et al.[7] The rest of the risk allele counts (for genes other than *CFH* and *ARMS2*) are available only through the Awh et al 2013 study and were used in this analysis.

Table 1. Counts of *CFH* Genotype by *CFH* Risk Allele Number in the 2 Data Sets

Data Set	*CFH* Risk Allele Number								
	0			*1*			*2*		
AREDS									
rs3766405\rs412852	CC	CT	TT	CC	CT	TT	CC	CT	TT
CC	0	243	34	0	0	0	536	0	0
CT	0	0	113	0	376	0	0	0	0
TT	0	0	111	0	0	0	0	0	0
Arctic									
rs3766405\rs412852	CC	CT	TT	CC	CT	TT	CC	CT	TT
CC	1	2	8	6	168	24	353	1	1
CT	1	1	35	2	255	51	0	4	0
NR	0	0	1	0	0	0	0	0	0
TT	1	1	69	0	0	4	0	0	0

AREDS = Age-Related Eye Disease Study; NR = Not Recorded.

Because we selected the overlapping patients (n = 879) based on the 2 published studies, we checked the balance in baseline covariates and genotypes. The balance assessment would help to examine whether this overlapping subset of patients is an unbiased sample of the AREDS cohort participating in the randomized controlled trial. In addition, the genotypic subgroups in Awh et al[3]—defined by the number of risk alleles of *CFH* and *ARMS2* genes—were not prespecified, but rather were identified from a forward stepwise variable selection procedure using the same outcome data that were used for their primary analysis.[9] Such a procedure is well known to increase the risk of false-positive results because of multiple testing. To avoid this problem, we scanned through the complete, unselected set of genetic markers, testing for possible interactions between the treatment and each marker, while explicitly accounting for the total number of hypotheses investigated. Specifically, to test for interactions, we used the Cox proportional hazards model for prediction of AMD progression as a function of treatment, a given genetic marker, and their interaction term. The interaction term addresses the hypothesis that the effects of treatment depend on the predictor. For instance, a statistically significant interaction between treatment and gender would mean that the benefit of treatment is different between men and women. Because there are 11 markers in total, the interaction between treatment and each gene was tested in a separate Cox proportional hazards model. We adjusted for multiple testing using the Bonferroni correction with a nominal family-wise error rate of 0.05; that is, we adjusted our results to be equivalent to if we had tested just a single hypothesis with a threshold of 0.05 for statistical significance.

To account for any imbalance between groups, we further tested for gene-by-treatment interactions by fitting an adjusted Cox proportional hazards model controlling for clinical variables for each gene. These clinical variables include age (1 degree of freedom [df]), BMI (1 df), gender (1 df), smoking status (2 df), baseline AMD category (1 df), and education (high school or not, 1 df). Bonferroni correction was used for this set of adjusted analyses with the same nominal family-wise error rate. Ties in time to progression were handled by Efron's method in all Cox regression models.[10] All statistical analyses were conducted using SAS software version 9.4 (SAS Institute, Cary, NC).

Memorial Sloan Kettering Cancer Center: Prediction of Treatment Response

Of the 2 members of the MSKCC team, one (A.J.V.) was aware that the goal of the analysis was to determine whether certain risk SNPs predicted treatment response, whereas the other (M.J.A.) was blinded to this purpose and acted as an independent biostatistician tasked with identifying any variables known at baseline that could predict the benefit of treatment.

In our analyses, we focused on 752 patients who were treated by antioxidants plus zinc (the currently recommended treatment) or placebo (as the comparison group) and for whom we had *CFH* status as defined by Awh et al using rs3766405 and rs412852 genotyping information, including those who were not analyzed by Awh et al but for whom we had the corresponding SNPs required to assess *CFH* status (n = 752). Clinical information, demographics, and *ARMS2* (c.372_815de1443ins54) data were available for 752 participants; rs1061170, rs1410996, and rs10490924 data were available for 601 patients. We created a series of Cox regression models predicting time to progression of AMD in terms of a baseline predictor, treatment status (antioxidants plus zinc vs. placebo), and the interaction between the predictor and treatment. For models including SNPs, we generated 2 dummy variables to represent SNP status and 2 dummy variables to represent the interaction terms and jointly tested the 2 interaction terms. Predictors of interest included demographic characteristics (age, smoking status, BMI, gender, and baseline AMD score) and genotypes from Chew et al (rs412852 [*CFH*], rs3766405 [*CFH*], rs1061170 [*CFH*], rs1410996 [*CFH*], and rs10490924 [*ARMS2*] status as defined by Chew et al). Additionally, we hypothesized that there is an interaction between treatment and the baseline risk of progression. To estimate baseline risk of progression, we used univariate Cox regression for patients in the placebo group to test for an association between baseline patient characteristics and progression among patients treated with placebo. We then used a backward selection procedure with a threshold of $P < 0.2$ to select predictors for inclusion in the risk model among the candidate predictors that were shown to be associated univariately with the outcome. Candidate predictors included age, smoking status, BMI, gender, baseline AMD severity, diabetes, high blood pressure, angina, cancer, arthritis, rs412852, rs3766405, rs1061170, rs1410996, and rs10490924. Statistical analyses were conducted using Stata software version 13 (Stata Corp., College Station, TX).

Results

MD Anderson Cancer Center: Data Concordance

The concordance at the level of genotype calls between the 2 data sets was good: concordance rates were 98.9% (869/879) for rs412853, 98.5% (866/879) for rs3766405, 97.6% (858/879) for both *CFH* SNPs, and 96.9% (852/879) for *ARMS2* indel calls. Although consistency between genotype call and risk allele counts was acceptable for AREDS data, in the Arctic data, not all samples with the same *CFH* genotypes had the same risk allele number (Table 1). For instance, of the 86 patients Arctic assigned *CFH* genotypes of rs3766405 = CT and rs412852 = TT, 35 were assigned a *CFH* risk allele number of 0 and 51 were assigned a *CFH* risk allele of 1.

Progression data also differed between groups. The results we obtained from applying our algorithm to the data from AREDS very closely matched the results they reported (we disagreed with 4/1413 times and 2/1413 status assignments). However, the outcome calls supplied disagreed with the raw data for 86 of the 989 samples examined by Arctic. Differences fell into 3 categories, including incorrect calls (45 cases), nonmonotonic progression patterns for some patients (e.g., going from 3 to 4 and then back to 3 [28 cases]), and incorrect but close follow-up times (13 cases). We were unable to identify an algorithm that would yield the data reported by Arctic.

In a Cox regression model of GTG 2 patients by treatment group using the samples examined by both AREDS and Arctic, the overall model was not significant (likelihood ratio statistic of 6.47 on 3 df; $P = 0.091$), although the results of a separate test for zinc alone were significant ($z = 2.25$; $P = 0.024$) before adjusting for multiple testing. These samples were part of the initial cohort examined by Awh et al in 2013. When we used the same approach to examine the GTG 2 patients examined only by AREDS, not only was the overall model not significant (likelihood ratio statistic of 0.51 on 3 df; $P = 0.9$), but also the effect of zinc alone was not significant ($z = 0.25$; $P = 0.8$). Hence, we were unable to replicate in an independent test set the strongest genotype−treatment interaction claimed.

We undertook further analysis to determine possible reasons for the contradictory findings. Differences in underlying data, such as for progression times or genotype calls, did not have a large impact, with small changes in odds ratios, such as from 2.14 to 2.08. Power similarly seems not to be an important issue: with 989 patients in the original analysis compared with 534 in our replication, the difference in the width of the CI is approximately 35%. Our results still would have been far from statistically significant had we obtained an identical central estimate but a 35% narrower CI. Hence, the primary reasons why our findings differed from those of Awh et al are overfit and multiple testing, which constitute the traditional rationale for independent replication.

Duke University: Interaction between Genotype and Treatment

Table S1 (available at www.aaojournal.org) summarizes the baseline clinical characteristics of the overlapping sample by treatment group, and Table S2 (available at www.aaojournal.org) presents the marker information by treatment group. The structure of these 2 tables is similar to that of Table 4 reported in Chew et al.[4] We did not find covariate or genotype imbalance across treatment groups except for a *P* value for 1 genotype of 0.006. This was for *C2*, an SNP that is not part of the genotype of purported value for treatment decision making. Because the probability of observing a *P* value of 0.006 or less when

Table 2. *P* Values for Testing Interaction Effects (Bonferroni-Corrected Significance Threshold = 0.0045)

Gene	Degrees of Freedom	Marker	*P* Value	*P* Value (Covariate Adjusted)
CFH	6	rs3766405	0.6	0.6
CFH	6	rs412852	0.072	0.018
C3	6	rs2230199	0.033	0.017
C2	3	rs4151669	1	0.8
CFB	3	rs522162	0.7	0.19
CFI	6	rs10033900	0.6	0.3
TIMP3	4	rs9621532	0.8	0.6
LPL	6	rs1268919	0.6	0.6
LIPC	3	rs492258	0.12	0.3
ABCA1	6	rs1883025	0.3	0.8
ARMS2	6	372_815del443ins54	0.058	0.089

Table 3. Treatment Effect Estimates and Bonferroni-Corrected Confidence Intervals by Number of Risk Alleles

Risk Alleles	Unadjusted Analysis		Covariate-Adjusted Analysis	
	Hazard Ratio	*Confidence Interval*	*Hazard Ratio*	*Confidence Interval*
CFH (rs412852)				
0	0.62	0.17–2.24	0.79	0.21–2.93
1	0.58	0.30–1.11	0.50	0.24–1.01
2	1.06	0.53–2.12	1.21	0.57–2.56
CFH (rs3766405)				
0	0.77	0.08–7.84	0.89	0.09–9.12
1	0.62	0.28–1.37	0.63	0.27–1.45
2	0.81	0.47–1.41	0.84	0.46–1.52
C3				
0	0.74	0.39–1.38	0.74	0.38–1.47
1	0.76	0.39–1.48	0.85	0.42–1.68
2	0.41	0.04–4.15	0.41	0.04–4.51
ARMS2 (372_815del443ins54)				
0	1.02	0.45–2.28	0.99	0.43–2.27
1	0.69	0.36–1.32	0.69	0.34–1.38
2	0.54	0.21–1.36	0.63	0.22–1.77

conducting 17 independent tests of 17 true null hypotheses is approximately 0.1, we concluded that the clinical values and genotypes of patients were distributed evenly across treatment groups. Thus, the overlapping sample could approximate a randomized study, and hence an unadjusted analysis should be unbiased.

For each marker, we fit an unadjusted Cox proportional hazards model including only treatment (3 df), genotype (assuming a codominant model if applicable), and their interactions. The covariate-adjusted Cox proportional hazards model also was used to assess further whether the conclusions change after taking into account the baseline patient characteristics. We used a Bonferroni-corrected significance threshold of 0.05/11 = 0.0045 to account for multiple testing. The *P* values for testing interaction effects in the Cox models are presented in Table 2. One interaction term is significant from the unadjusted analysis without accounting for multiplicity, *C3*, which is not the genotype claimed to be of value by Awh et al. Possible interaction was suggested between *CFH* (rs412852) and treatment from an adjusted analysis, but before multiplicity adjustment. No interaction term was found to be significant after controlling for multiple testing.

Table 4. Tests of Interaction between Treatment and Patient Characteristics

Patient Characteristic	*P* Value
Age (n = 794)	0.9
Smoking status (current and former vs. never; n = 794)	0.3
BMI (n = 794)	0.5
Baseline AMD score (3a and 3b vs. 4a and 4b; n = 794)	0.3
Gender (n = 794)	0.4
rs3766405 (n = 752)	0.5
rs412852 (n = 752)	0.059
rs1061170 (n = 601)	0.069
rs1410996 (n = 601)	0.15
rs10490924 (n = 601)	0.013
ARMS2 (c.372_815del443ins54; n = 752)	0.5
CFH status (Awh et al definition; n = 752)	0.057

AMD = age-related macular degeneration; BMI = body mass index. Each patient characteristic was tested in a separate Cox model.

Although we did not find a treatment–genotype interaction, we nonetheless estimated treatment effects by subgroup, addressing the hypothetical of whether treatment could be harmful in a subgroup. Table 3 presents the central estimates of treatment effect by genotype subgroup for the comparison of antioxidant plus supplement (results for all treatments are shown in Tables S3 and S4, available at www.aaojournal.org). In no subgroup was there any evidence to support a large increase in risk from antioxidants and zinc, with the highest hazard ratio being 1.06. For *C3*, the only gene whose genotype showed a conventionally significant interaction in the unadjusted analysis, hazard ratios for all genotype subgroups were well below 1, suggesting benefit irrespective of genotype.

Memorial Sloan Kettering Cancer Center: Prediction of Treatment Response

We did not find sufficient evidence of an interaction between treatment and any patient characteristics of interest except corresponding to the interaction with SNP rs10490924 status (Table 4). Upon further investigation, the interaction of treatment and rs10490924 status of GG versus GT and TT was significant, whereas the interaction with GT versus GG and TT was not ($P = 0.044$ and $P = 0.5$, respectively). Table 5 represents patient risk of progression within 5 years by rs10490924 and Awh et al–defined *CFH* status. Among all patients, treatment with antioxidants plus zinc was shown to be associated with decreased risk of progression at 5 years (6.9%; 95% CI, 0.8%–13%). Among the 41% of participants with an rs10490924 status of GG, there was not sufficient evidence of a difference in the risk of progression by treatment; however, among patients with an rs10490924 status of TT or GT, antioxidants plus zinc treatment was associated with a decreased risk of progression at 5 years of 13%. Although the interaction between Awh et al–defined *CFH* risk copies and treatment did not meet conventional levels of statistical significance, we nonetheless examined the patients' risk of progression within 5 years by *CFH* status. Among the 36% of patients with 2 *CFH* risk copies, there was no evidence of a difference in the risk of progression by treatment; however, among most patients, the treatment was shown to be associated with a decreased risk of progression at 5 years of 12%.

Table 5. Five-Year Kaplan-Meier Estimates of Risk of Progression by Treatment Arm and Genotype Status

Single Nucleotide Polymorphism Status	Placebo (%)	Treatment (%)	Difference (% [95% Confidence Interval])
All patients (n = 794)	29	22	6.9 (0.8–13)
rs10490924 TT or GT (n = 357)	34	21	13 (4.0–23)
rs10490924 GG (n = 244)	14	17	−2.5 (−12 to 6.6)
CFH 0 or 1 risk copies (n = 484)	30	18	12 (4.5–20)
CFH 2 risk copies (n = 268)	28	29	−1.0 (−12 to 9.8)

A positive value represents a treatment benefit. *CFH* risk copies were determined based on the definition by Awh et al.[3,8]

Variables selected for inclusion in the multivariable model included smoking status (current vs. former and never), age, baseline AMD status (3a and 3b vs. 4a and 4b), and rs10490924 and rs1410996 status. The interaction between treatment and the estimated risk of progression had a patient received placebo was significant ($P = 0.032$; n = 601), with a larger improvement in PFS among patients at higher baseline risk.

We did not find sufficient evidence to suggest that the CFH status defined by Awh et al influenced the effectiveness of the treatment. Although there was a significant interaction between treatment and rs10490924, this single significant *P* value is not compelling in the context of multiple testing: even if we ignore the analyses of nongenotype predictors, a Bonferroni-adjusted *P* value would be 0.091.

Discussion

Our 3 different statistical groups conducted independent but complementary analyses to determine whether genotyping of *CFH* and *ARMS2* should be used to guide the decision of whether to use antioxidants and zinc to prevent AMD progression. All 3 groups concluded that genotyping is unwarranted. The MD Anderson Cancer Center team found important errors of summarization in the data set used in the original paper by Awh et al[3,8] supporting genotyping. Moreover, no evidence ($P = 0.9$) was found for a key claim of Awh et al when tested on independent samples. The Duke group analyzed all 11 of the genotypes examined by Awh et al. There were no statistically significant interactions after adjusting for multiple testing. Memorial Sloan Kettering Cancer Center took a blank slate approach to predicting treatment response. Although there was evidence that, in general, patients at higher baseline risk had a greater improvement in PFS with supplements, the evidence did not support genotyping. There were no statistically significant interactions after adjusting for multiple testing and there was no support for the critical claim that risk of AMD progression is much higher patients with some genotypes. A key consideration here is that the treatment—antioxidants and zinc—is benign, and so there is a greater burden of proof to demonstrate a poorer outcome on the primary end point for anyone advocating a test to predict treatment response.

Several investigators previously have attempted to resolve the discrepancy between the studies by Awh et al and Chew et al. In a sophisticated analysis of AREDS data using the eye, rather than the patient, as the level of analysis, Seddon et al[11] reported that supplementation was effective only for the subgroup of patients with the TT genotype for *CFH Y402H* or *ARMS2*. The authors concluded: "The effectiveness of antioxidant and zinc supplementation appears to differ by genotype." We find the results of Seddon et al actually very comparable with our own and believe that their conclusions are not supported by the results they present. The *P* values for the interaction between treatment and genotype for the main end point of advanced AMD are 0.069 for *CFH Y402H* and 0.024 for *ARMS2*. These are not impressive *P* values given that 4 hypotheses were tested, even leaving aside that *CFH Y402H* or *ARMS2* were selected from a total of 11 genes examined by Awh et al. Furthermore, just as in the current analysis, Seddon et al did not find harm associated with treatment. The highest central estimate of hazard ratio in any subgroup was 1.04, casting doubt on the claim by Awh et al[3,8] of a "deleterious interaction" sufficient to justify a genomic test. Indeed, one could return to the original study by Awh et al to make the same point about multiple testing: the lowest *P* value reported for any interaction term was 0.01. This is arguably significant only if we ignore that *CFH* and *ARMS2* first were selected from 11 genes.

That is, our conclusions differ from those of prior authors at least in part because of this issue of multiple testing. We believe multiple testing to be a fundamental and uncontroversial aspect of statistical methodology. As a simple illustration, if a man were to flip a coin 1000 times per day for 1 month, the probability that the final proportion of heads is statistically different from 50% is, as expected, 0.05. However, there is approximately an 80% chance that he will throw statistically significantly more or fewer heads on at least 1 day. We also might point out that the lowest *P* value in any analysis, and one lower than the *P* value for any interaction term reported in any study, is for baseline treatment differences between groups in the prevalence of rs4151669 (Table S2, available at www.aaojournal.org), a finding almost impossible to explain in terms other than chance. The issue of multiple testing was identified as a concern by Chew et al, by Wittes and Musch[6] in their editorial, and independently by each of the 3 groups in the current analysis. It is well known that if multiple testing is ignored, discoveries are unlikely to be replicable in data sets other than those used to discover them in the first place. Indeed, much contemporary methodologic work on appropriate statistical analysis of genetic data focuses exclusively on multiple testing. Our concern about multiple testing is borne out by the analysis conducted by the MD Anderson Cancer Center group on an independent group of samples, which found no treatment−genotype interaction.

There are several differences in results among our 3 groups. Both Duke University and MSKCC calculated interaction terms between treatment and genotype. The former investigated the genotypes in the Awh et al study and the latter those in the Chew et al study. The *P* values for the SNPs investigated by both groups are not entirely consistent because MSKCC looked only at the antioxidant plus zinc group, whereas Duke University examined interactions across all 4 treatments. That said, the qualitative conclusion reached by all 3 of our groups—that of no significant interaction after correction for multiple testing—is the same.

Our findings illustrate the importance of replication for marker studies. One of us (A.J.V.) has been involved in the development of a diagnostic test. This was tested in 9 separate retrospective studies including more than 15 000 patients before a prospective validation study was conducted.[12] It was only after this final prospective study that the test was made commercially available. The problem with the AMD genotyping controversy may be

related to the fact that all claims about the need for genotyping were based on retrospective analysis of a single study, and the test was commercialized without prospective replication on additional empirical data.

We cannot prove a negative. It may well be that, with further data collection and analysis, pharmacogenetic markers will be developed that can guide chemoprevention of AMD. It also may be the case that further research on *ARMS2*, *CFH*, or other SNPs included in this analysis demonstrate their value for treatment-related decision making. Our claim is that, at the current time, we do not have good reason to believe that genotyping will do more good than harm.

In conclusion, our separate statistical groups analyzed data from the AREDS study using 3 separate but complementary statistical approaches. We found no evidence to support the use of genotyping to inform chemoprevention of AMD. Patients who meet current criteria for supplementation—extensive intermediate-size drusen, at least 1 large druse, noncentral geographic atrophy in 1 or both eyes, or advanced AMD or vision loss resulting from AMD in 1 eye—and who have no contraindications to supplements, such as smoking, should be offered zinc and antioxidants without consideration of genotype.

References

1. Age-Related Eye Disease Study Research Group. 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*. 2001;119(10):1417-1436.
2. Klein ML, Francis PJ, Rosner B, et al. *CFH* and *LOC387715/ARMS2* genotypes and treatment with antioxidants and zinc for age-related macular degeneration. *Ophthalmology*. 2008;115(6):1019-1025.
3. Awh CC, Lane AM, Hawken S, et al. *CFH* and *ARMS2* genetic polymorphisms predict response to antioxidants and zinc in patients with age-related macular degeneration. *Ophthalmology*. 2013;120(11):2317-2323.
4. Chew EY, Klein ML, Clemons TE, et al. No clinically significant association between *CFH* and *ARMS2* genotypes and response to nutritional supplements: AREDS report number 38. *Ophthalmology*. 2014;121(11):2173-2180.
5. Awh CC, Zanke BW. Re: Chew et al.: no clinically significant association between *CFH* and *ARMS2* genotypes and response to nutritional supplements: AREDS report number 38 (*Ophthalmology*. 2014;121:2173−2180). *Ophthalmology*. 2015;122(8):e46.
6. Wittes J, Musch DC. Should we test for genotype in deciding on age-related eye disease study supplementation? *Ophthalmology*. 2015;122(1):3-5.
7. Chew EY, Klein ML, Clemons TE, et al. Author reply to PMID 24974817. *Ophthalmology*. 2015;122(8):e46-47.
8. Awh CC, Hawken S, Zanke BW. Treatment response to antioxidants and zinc based on *CFH* and *ARMS2* genetic risk allele number in the Age-Related Eye Disease Study. *Ophthalmology*. 2015;122(1):162-169.
9. Chew EY, Klein ML, Clemons TE, et al. Genetic testing in persons with age-related macular degeneration and the use of the AREDS supplements: to test or not to test? *Ophthalmology*. 2015;122(1):212-215.
10. Efron B. The efficiency of Cox's likelihood function for censored data. *J Am Statist Assoc*. 1977;72(359):557-565.
11. Seddon JM, Silver RE, Rosner B. Response to AREDS supplements according to genetic factors: survival analysis approach using the eye as the unit of analysis. *Br J Ophthalmol*. 2016;100(12):1731-1737.
12. Parekh DJ, Punnen S, Sjoberg DD, et al. A multi-institutional prospective trial in the USA confirms that the 4Kscore accurately identifies men with high-grade prostate cancer. *Eur Urol*. 2015;68(3):464-470.

Footnotes and Financial Disclosures

Originally received: July 24, 2017.
Final revision: September 1, 2017.
Accepted: September 8, 2017.
Available online: October 9, 2017. Manuscript no. 2017-1688.

[1] Department of Epidemiology and Biostatistics, Memorial Sloan Kettering Cancer Center, New York, New York.

[2] Department of Biostatistics and Bioinformatics, Duke University, Durham, North Carolina.

[3] Department of Bioinformatics and Computational Biology, MD Anderson Cancer Center, Houston, Texas.

*These authors contributed equally as first authors.

‡These authors contributed equally as senior authors.

Financial Disclosure(s):
The author(s) have made the following disclosure(s): A.J.V.: Supported in part by the National Cancer Institute, National Institutes of Health, Bethesda, Maryland (Cancer Center Support grant to MSKCC, no.: P30-CA008748).

No funds were received directly for analyses of this project.

HUMAN SUBJECTS: No human subjects were included in this study.

Author Contributions:

Conception and design: Assel, Li, Wang, Allen, Baggerly, Vickers

Analysis and interpretation: Assel, Li, Wang, Allen, Baggerly, Vickers

Data collection: Assel, Li, Wang, Allen, Baggerly, Vickers

Obtained funding: none

Overall responsibility: Assel, Li, Wang, Allen, Baggerly, Vickers

Abbreviations and Acronyms:
AMD = age-related macular degeneration; **AREDS** = Age-Related Eye Disease Study; ***ARMS2*** = age-related maculopathy susceptibility 2; **BMI** = body mass index; ***CFH*** = complement factor H; **CI** = confidence interval; **df** = degree of freedom; **GTG** = genotype group; **indel** = insertion/deletion; **MSKCC** = Memorial Sloan Kettering Cancer Center; **NIH** = National Institutes of Health; **PFS** = progression-free survival; **SNP** = single nucleotide polymorphism.

Correspondence:
Andrew J. Vickers, PhD, Epidemiology and Biostatistics, Memorial Sloan Kettering Cancer Center, 485 Lexington Avenue, 2nd Floor, New York, NY 10017. E-mail: vickersa@mskcc.org.