Published in final edited form as:

Ophthalmology. 2018 March ; 125(3): 391–397. doi:10.1016/j.ophtha.2017.09.008.

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. Assel¹, Fan Li², Ying Wang³, Andrew S. Allen², Keith A. Baggerly³, and Andrew J. Vickers¹

¹Memorial Sloan Kettering Cancer Center, New York, New York

²Duke University, Durham, North Carolina

³MD Anderson Cancer Center, Houston, Texas

Abstract

PURPOSE—Recent years have seen considerable controversy as to whether genotyping should be part of standard care for patients with age-related macular degeneration (AMD) 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) trial, 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 category > 1 and genotyping data available. There is imperfect overlap between data from the two groups with respect to measurements made: the largest common set involves 879 participants for which the same CFH and ARMS2 SNPs 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 higher risk patients had more to gain from treatment, we were unable to replicate any genotype treatment interactions after adjusting for multiple testing. We tested one genotype claim on an independent set of data, with negative results. Even if

MJA, FL and YW wish to be recognized as joint first authors; ASA, KAB and AJV wish to be recognized as joint senior authors

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

we assumed that interactions did in fact 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

Introduction

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% C.I. 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. published a pharmacogenomic paper suggesting that the effects of antioxidants and zinc on AMD in the AREDS trial might be influenced by genotype², 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 CC genotype for CFH Y402H (44% vs. 39%) compared to 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 appeared 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., to attempt a replication⁴. 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 tests for interaction being non-significant. 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 Awh⁵, who claimed that the Chew et al. paper refutes any claim of overall benefit for supplementation and that a separate editorial, written by a well known statistician/epidemiologist team (Janet Wittes and David Musch)⁶, supported the

genotyping. In response, Chew et al. claimed that Awh and Zanke had misinterpreted their paper and that, in fact, the Wittes and Musch editorial favored their own position.⁷

To help resolve this debate, the Office of Intramural Research at the National Institutes of Health (NIH) asked our three biostatistical groups to independently reexamine the data used by Awh et al. and Chew et al. in order to determine whether genotyping should be part of the clinical decision whether to use supplements for AMD prevention. Here 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 two 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 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 (dbGaP), and SNP/sequencing data is now available for an ever-increasing subset, though much less was available when the debate began. For their studies^{3, 8}, Awh et al. 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 indel status for ARMS2 at one location. Chew et al.^{4, 9} looked at data from 1237 patients for which they had CFH and ARMS2 genotype data at SNPs other than those used by Awh et al. (rs1061170 and rs1410996 for CFH, rs10490924 for ARMS2; summarized in their Figure 1b), and at data from 1413 patients measured using exactly the same SNPs used by Awh et al. (summarized in their Figure 1c). In all, genotype data from these three Awh et al. locations is available for 1523 participants: 879 were measured by both groups, 110 were only measured by Awh et al, and 534 were only measured by Chew et al. All data can be matched using anonymized AREDS patient identifiers.

The genotype data for patients measured at the above three SNPs underwent several levels of summarization. First, there were the raw genotype assessments (AA, AB, or BB) at each of the three 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 one 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 two genes are used to assign patients to genotype groups (GTGs). Proposed treatment differentiation would occur at the GTG level.

The three statistical groups decided to work independently on three separate approaches to the replication problem. The MD Anderson group focused primarily on data checking and

evaluating concordance between different data sets. Duke's role was to replicate the key findings of Awh et al. concerning interactions between genotype and outcome. Memorial Sloan Kettering Cancer Center (MSKCC) 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: Data concordance

We received raw data on patients from the AREDS trial¹ linking times to AMD disease progression to CFH and ARMS2 genotypes and treatment group, from both Awh et al. ("Arctic")^{3, 8} and AREDS investigators^{4, 9}. The data also contained various clinical covariates such as age, sex, race, BMI and smoking history.

Since unappreciated differences between data sets could explain some of the published inconsistencies, we first 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 two 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 one eye rated as category 4 at the outset.

We split the data into three 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 three data sets contains AREDS ID, CFH genotype, ARMS2 genotype, treatment group index, progression status and time from each of the two groups. The cleaned data were presented to and approved by all three groups before they started their own independent analyses.

As noted in Chew et al. 2014⁴, the longer follow-up times now available include times after the end of randomization for the initial trial, at which point the different treatment groups were all shifted to receive the AREDS formulation. Since this could distort treatment differences, we chose to work with the PFS (progression-free survival) 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 genotype groups (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 datasets examined just by one group. For this, we focused on GTG2 (both CFH SNPs are "CC", and ARMS2 category is "11") from Awh et al. 2015⁸, as this was the subgroup for which the strongest claims were made. We started with the 120 GTG2 patients examined by both Arctic and AREDS investigators. Then as a validation test, we performed the same analyses using the 75 GTG2 patients examined by AREDS group alone. This is similar to the approach taken in Chew et al. 2015⁹, but uses the direct matching we were able to obtain with access to datasets from both groups, that Chew et al. 2015 did not have. Analyses were conducted using R-2.3.0.

Duke: Interaction between genotype and treatment

We include in the analysis 879 AREDS patients that have a high risk of progression to late AMD and that are contained 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 non-compliance or crossover effects in the observational phase of study may introduce bias, we confined our analysis to the period of the randomized controlled trial in 2001.

A comprehensive set of 11 genetic markers for AMD has been identified by Awh et al. (see Table 1³), which 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%). Since the process of summarizing CFH genotypes (rs3766405, rs412852) into risk allele counts was more consistently applied 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_815del443ins54) were highly concordant between the two 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 only available through the Awh et al. 2013 study and were used in this analysis.

Since we selected the overlapping patients (n = 879) based on the published two studies, we checked balance in baseline covariates and genotypes. The balance assessment would help 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 pre-specified, but rather identified from a forward stepwise variable selection procedure using the same outcome data that were used for their primary analysis⁹. Such a procedure is well known to increase the risk of false positives due to 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 the 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, say, gender, would mean that the benefit of treatment is different between men and women. Since 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 adjusted Cox proportional hazards model controlling for clinical variables for each gene. These clinical variables include age (1 degree of freedom, df), body mass index (1 df), gender (1 df), smoking status (2 df), baseline AMD category (1 df) and education (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 models¹⁰. All statistical analyses were conducted using SAS v 9.4 (Cary, NC).

MSKCC: prediction of treatment response

Of the two members of the MSKCC team, one (AV) was aware that the goal of the analysis was to determine whether certain risk SNPs predicted treatment response whereas the other (MA) 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 which we had the corresponding SNPs required to assess CFH status (n=752). Clinical information, demographics and ARMS2 (c.372_815del443ins54) 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 two dummy variables to represent SNP status and two dummy variables to represent the interaction terms and jointly tested the two 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), 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. In order to estimate baseline risk of progression, we used univariable Cox regression on 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 backwards 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 univariately associated 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 13 (Stata Corp., College Station, TX).

Results

MD Anderson: Data concordance

The concordance at the level of genotype calls between the two 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. While 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 three categories including: incorrect calls (45 cases); non-monotonic progression patterns for some patients (e.g., going from 3 to 4 and then back to 3: 28 cases); incorrect but close follow-up times (13 cases). We were unable to identify an algorithm that would yield the data reported by Arctic.

In Cox regression model of GTG2 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, $p = 0.091$), though the separate test for zinc alone is 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. 2013. When we used the same approach to examine the GTG2 patients examined only by AREDS, not only was the overall model not significant (likelihood ratio statistic of 0.51 on 3df, $p = 0.9$), but 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 appears not to be an important issue: with 989 patients in the original analysis compared to 534 in our replication, the difference in the width of the confidence interval is about 35%. Our results would still have been far from statistically significant had we obtained an identical central estimate but a 35% narrower confidence interval. 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: Interaction between genotype and treatment

Supplementary sTable 1 summarizes the baseline clinical characteristics of the overlapping sample by treatment group, and supplementary sTable 2 presents the marker information by treatment group. The structure of these two tables is similar to the 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 one genotype of 0.006. This was for C2, a SNP that is not part of the genotype of purported value for treatment decision-making. Since the probability of observing a p value 0.006 when conducting 17 independent tests of 17 true null hypotheses is approximately 0.1, we conclude that the clinical values and genotypes of patients are evenly distributed 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 (3df), genotype (assuming a co-dominant model if applicable) and their interactions. The covariate-adjusted Cox proportional hazards model was also used to assess further whether the conclusions change after taking into account the baseline patient characteristics. We used a Bonferroni-corrected significance threshold of $0.05 \div 11 = 0.0045$ 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 further suggested between the CFH (rs412852) and treatment from an adjusted analysis, but before multiplicity adjustment. No interaction term is found significant after controlling for multiple testing.

Even though 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 supplemental tables s3 and s4). In no subgroup is 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 shows a conventionally significant interaction in the unadjusted analysis, hazard ratios for all genotype subgroups are well below 1, suggesting benefit irrespective of genotype.

MSKCC: 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). On 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 0.5 , respectively). Table 5 represents the patients' risk of progression within 5 years by rs10490924 and Awh-defined CFH status. Among all patients, treatment with antioxidants plus zinc was shown to be associated with decreased risk of progression at 5 years of 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 was associated with a decreased risk of progression at 5 years of 13%. Although the interaction between Awh-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 two CFH risk copies there was no evidence of a difference in the risk of progression by treatment,

however, among the majority of patients the treatment was shown to be associated with a decreased risk of progression at 5 years of 12%.

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), 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 progression-free survival among patients at higher baseline risk.

We did not find sufficient evidence to suggest that the CFH status defined by Awh 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 non-genotype predictors, a Bonferroni adjusted p value would be 0.091.

Discussion

Our three 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 progression of AMD. All three groups concluded that genotyping is unwarranted. The MD Anderson team found important errors of summarization in the data set used in the original paper, Awh et al., 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 in Awh et al. There were no statistically significant interactions after adjusting for multiple testing. MSKCC took a “blank slate” approach to predicting treatment response. Although there was evidence that, in general, patients at higher baseline risks had a larger improvement in progression-free survival with supplements, the evidence did not support genotyping. There were no statistically significant interactions after adjusting for multiple testing and 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 endpoint for anyone advocating a test to predict treatment response.

Several investigators have previously attempted to resolve the discrepancy between the Awh et al. and Chew et al. papers. 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 only effective for the subgroup of patients with 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 to our own and believe that their conclusion is not supported by the results they present. The p-values for the interaction between treatment and genotype for the main endpoint of advanced AMD are 0.069 for *CFH Y402H* and 0.024 for *ARMS2*. These are not impressive p-values given that four 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 Awh et al. claim of a “deleterious interaction” sufficient to justify a genomic test. Indeed, one could return to the original Awh et al. paper 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* were first selected from 11 genes.

In other words, our conclusions differ from 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 a day for a month, the probability that the final proportion of heads is statistically different from 50% is, as expected, 0.05. However, there is about an 80% chance that he will throw statistically significant more or fewer heads on at least one day. We might also point out that the lowest p-value in any analysis, and one lower than the p-value for any interaction term reported in any paper, is for *baseline* treatment differences between groups in the prevalence of rs4151669 (see supplementary table sTable 2) 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 in their editorial and independently by each of the three groups in the current analysis. It is well known that if multiple testing is ignored, discoveries are unlikely to be replicable in datasets other than those used to discover them in the first place. Indeed, much contemporary methodological work on appropriate statistical analysis of genetic data exclusively focuses on multiple testing. Our concern about multiple testing is borne out by the analysis conducted by the MD Anderson group on an independent group of samples, which found no treatment genotype interaction.

There are several differences in results between our three groups. Both Duke and MSKCC calculated interaction terms between treatment and genotype. The former investigated the genotypes in the Awh et al. paper and the latter those in Chew et al. 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 examined interactions across all four treatments. That said, the qualitative conclusion reached by all three 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 (AV), has been involved in the development of a diagnostic test. This was tested in 9 separate retrospective studies including over 15,000 patients before a prospective validation study was conducted¹². 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, with commercialization of the test occurring 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 may also be the case that further research on *ARMS2*, *CFH* or other SNPs included in this analysis demonstrate their value for treatment 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 three 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 due to AMD in 1 eye - and who have no contraindications to supplements, such as smoking, should be offered zinc and antioxidants without consideration of genotype.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding acknowledgements.

No funds were received directly for analyses of this project. Andrew J. Vickers is supported in part by a National Institutes of Health/National Cancer Institute Cancer Center Support Grant to MSKCC (grant number P30-CA008748);

References

1. 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):1417–36. [PubMed: 11594942]
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):1019–25. [PubMed: 18423869]
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):2317–23. [PubMed: 23972322]
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):2173–80. [PubMed: 24974817]
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;121:2173–80). *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. [PubMed: 25542537]
7. Chew EY, Klein ML, Clemons TE, et al. Author reply: To PMID 24974817. *Ophthalmology*. 2015; 122(8):e46–7.
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):162–9. [PubMed: 25200399]
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):212–5. [PubMed: 25456150]
10. Efron B. The Efficiency of Cox's Likelihood Function for Censored Data. *Journal of the American Statistical Association*. 1977; 72(359):557–65.

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): 1731–7. [PubMed: 27471039]
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): 464–70. [PubMed: 25454615]

Highlights

Three separate statistical teams investigated whether genotype predicts response to supplements. All teams conclude that patients at risk of age-related macular degeneration progression should be offered zinc and antioxidants regardless of genotype.

Table 1

Counts of CFH genotype by CFH risk allele number in the two data sets

		CFH risk allele number=0			CFH risk allele number=1			CFH risk allele number=2		
AREDS Data										
rs3766405		CC	CT	TT	CC	CT	TT	CC	CT	TT
rs412852										
CC	0	243	34	0	0	0	0	536	0	0
CT	0	0	113	0	376	0	0	0	0	0
TT	0	0	111	0	0	0	0	0	0	0
Arctic Data										
rs3766405		CC	CT	TT	CC	CT	TT	CC	CT	TT
rs412852										
CC	1	2	8	6	168	24	353	1	1	1
CT	1	1	35	2	255	51	0	4	0	0
NR	0	0	1	0	0	0	0	0	0	0
TT	1	1	69	0	0	4	0	0	0	0

Table 2

P-values for testing interaction effects (Bonferroni-corrected significance threshold = 0.0045).

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

Table 4
Tests of interaction between treatment and patient characteristics

Each patient characteristic was tested in a separate Cox model.

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_815del1443ins54) (n=752)	0.5
CFH Status (Awh <i>et al.</i> definition) (n=752)	0.057

Table 5
5-year Kaplan-Meier estimates of risk of progression by treatment arm and genotype status

A positive value represents a treatment benefit. CFH risk copies were determined based on the definition in Awh *et al.*

SNP Status	Placebo	Treatment	Difference (95% CI)
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%, 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%, 9.8%)