Should We Test for Genotype in Deciding on Age-Related Eye Disease Study Supplementation?

Janet Wittes, PhD - Washington, DC
David C. Musch, PhD, MPH - Ann Arbor, Michigan

When we, a 71-year-old white woman and a 60-year-old white man, both nonsmokers, have our eyes examined, we will want to know whether we classify as Age-Related Eye Disease Study (AREDS) category 3, and therefore would likely benefit from taking the AREDS supplement, which contains high-dose antioxidants (β-carotene, vitamin C, and vitamin E) plus high-dose zinc. As a statistician (J.W.) who chaired the AREDS and AREDS2 Data Safety Monitoring Committee and an epidemiologist (D.C.M.) who recently wrote an editorial about the results of the AREDS trial,1 we are interested in the interchange between Awh et al2,3 (see article on page 162) and Chew et al4,5 (see article on page 212) not only professionally but also personally. If we indeed are at risk for development of neovascular age-related macular degeneration (AMD), we want to know whether we should follow the recommendations of Awh et al3 and undergo genetic testing to determine what supplements we should take or whether we should follow the guidance of Chew et al5 and take AREDS supplements without engaging in genetic testing.

We rely on 5 articles to inform our decision. All use data from the AREDS, a randomized, masked, clinical trial with 4757 participants assigned to 1 of 4 treatment groups: antioxidants and zinc placebo, zinc and antioxidant placebo, both antioxidants and zinc, or double placebo. First, Klein et al6 showed data from the AREDS suggesting that genetic variants may have differential response to treatments with antioxidants and zinc; however, the authors concluded that all variants of the genotypes they studied showed some benefit of AREDS supplements. Therefore, they did not recommend genetic screening. Their analysis was based on genetic data from 876 white participants of the AREDS trial who at baseline had intermediate AMD or unilateral advanced AMD.

Five years later, Awh et al2 analyzed AREDS data but reached different conclusions. These authors studied 995 white AREDS participants who had category 3 AMD in 1 eye and category 1 to 4 AMD in the fellow eye, and for whom peripheral blood-derived DNA was available. The authors considered 9 sets of risk alleles defined by complement factor H and age-related maculopathy sensitivity 2. They concluded that the best supplement (antioxidants, zinc, both, or neither) depends on the individual’s genotype (Table 1).

In an article countering those findings, Chew et al4 tested for evidence of a difference in the effect of treatment by genotype. The interaction P value for testing whether people’s response to the AREDS supplement differs by genotype was not significant when they looked at the genotypes described by Klein et al6 (P = 0.06) or those described by Awh et al2 (P = 0.45). Therefore, they concluded that AREDS supplementation reduces the rate of progression of AMD across all genotype groups.

Awh et al3 disputed the conclusions of Chew et al4, criticizing their division of the AREDS population. Moreover, Awh et al3 presented data that the authors claimed show clinically important differences by genotype in the effect of antioxidants and zinc.

Now, Chew et al5 have responded, criticizing Awh et al3 for presenting results that are biased and therefore difficult to interpret. The most important contribution of the article by Chew et al5 is the use of newly available genetic data from the AREDS. In applying grouping of genotypes in the study by Awh et al,3 Chew et al5 found that every group shows benefit of AREDS supplements.

In sorting out these conflicting articles, we are not going to dwell on some issues we have with them. We, as do Awh et al,3 find the presentation by Chew et al5 of 27 comparisons, many with small sample sizes, not clinically interpretable. Our more important agreement is with Chew et al5, who point out that Awh et al3 do not correct adequately for multiple statistical testing and use post hoc exploration of subgroups. Correcting for multiplicity is necessary in dealing with subgroups — shooting many arrows at a target improves the chances for even a poor archer; however, it is not clear how to deal with multiplicity in this case because Awh et al5 did not report how many subgroups were examined or which genotypes, if any, were prespecified and which were post hoc. Post hoc exploration of subgroups is treacherous because one never knows whether the results are due to inadvertent selection bias or if they are “true.”8 Many people search post hoc for subgroup effects. Such “data dredging” rewards the searcher with a high probability of finding differences between subgroups, but the probability is high that the finding is spurious. Replication is the only reliable way of assessing the validity of a post hoc finding.

Awh et al3 develop regression models to identify predictive variables; however, retrospective fitting of data is, at best, hypothesis-generating; the real test of a post hoc exploratory analysis is to find an independent data set to

Should our ophthalmologist send us for genotyping, we will use the recommendation as an opportunity for warning about the inferential perils of subgroup analyses…
serve to validate the finding. As noted by Awh et al, “Given the absence (to our knowledge) of an existing dataset appropriate for validation studies, confirmation of our findings in a different cohort is unlikely in the near future.”

Fortunately, such a dataset has emerged from the AREDS itself. Awh et al used a dataset of 995 white subjects with AREDS category 3 disease in at least 1 eye who had available DNA. Chew et al had access to an additional 526 patients from the AREDS with the same qualifications whose DNA has recently become available. The demographics of the test dataset (Awh et al) and the validation set (Chew et al) are nearly the same. Chew et al looked at the 4 genetic groups developed by Awh et al. Had the validation sample by Chew et al replicated the findings by Awh et al, then despite all the potential biases inherent in the analyses by Awh et al, we would have considered the recommendation to have genetic testing reasonable. In contrast to the finding by Awh et al, however, Chew et al found that the AREDS formulation (antioxidants plus zinc) showed benefit in delaying the development of neovascular AMD in all 4 genetic subgroups.

To understand what happened, consider Table 1. The first 3 columns and the fifth column are from Table 6 in the article by Awh et al, the fourth column is calculated, and the last column presents the genetic groups assigned by Awh et al.

Awh et al took the 9 genotypes identified in their earlier article, defined 4 categories on the basis of outcome (Table 1, last column), and fit statistical models that identify which supplements (antioxidants, zinc, or both) each of the 4 groups should use. It does not take a statistician or epidemiologist to forecast that the model will predict the results shown in Table 2. Yusuf et al point to subgroups defined by post-randomization variables as “improper” and not amenable to unbiased statistical testing. We distinguish between subgroups identified post hoc, but defined on the basis of baseline (i.e., pre-randomization) variables from the more problematic improper subgroups, whether defined a priori or post hoc, that are based on post-randomization variables. The genetic subgroups in the report by Awh et al are both post hoc and improper. The approach that Awh et al used in defining their genotype subgroups is circular: After selecting post-randomization outcomes observed by Awh et al to define subgroups in the study by Awh et al, they then tested those subgroups with the very same data.

The theoretical physicist Eugene Wigner describes a class Einstein taught in 1928. “He told us once: ‘Life is finite. Time is infinite. The probability that I am alive today is zero. In spite of this, I am now alive. Now how is that?’ None of his students had an answer. After a pause, Einstein said, ‘Well, after the fact, one should not ask for probabilities.’” Calculating probabilities after the fact is exactly what Awh et al did. Awh et al identified 9 genotypes defined by 0, 1, or 2 risk alleles in each of 2 genes. After the fact, they performed modeling, looking for \( P \) values, which are probabilities. Awh et al raised some hypotheses about genotype and treatment with the AREDS supplement. The validation study by Chew et al showed no evidence that those hypotheses were true. Thus, this statistician and this epidemiologist sent us for genotyping, we will use the recommendation as an opportunity for warning about the inferential perils of subgroup analyses, especially those that (similar to the study by Awh et al) are based on outcomes.

References


Footnotes and Financial Disclosures

Financial Disclosure(s): The author(s) have made the following disclosure(s): J.W.: Chaired — Data and Safety Monitoring Committees for the AREDS and AREDS2 trials. D.C.M.: Honoraria — Data and Safety Monitoring Committee participation for the AREDS2 sirolimus clinical trial from EMMES Corp.; Data Monitoring Committee participation for the clinical trial of emixustat HCL treatment for geographic atrophy associated with AMD from Acucela Inc.; Data and Safety Monitoring Board participation for the clinical trial of radiation treatment for neovascular AMD from Oraya Therapeutics Inc.