Age-Related Eye Disease Study Caveats

It is wonderful to finally see the data from the Age-Related Eye Disease Study (AREDS) addressing the issue of nutritional supplements in age-related macular degeneration (AMD). Predictably, no sooner were the results released than a commercial preparation was also released. We are concerned about the misplaced enthusiasm that has met the study findings and would like to advise a more temperate response.

We are only beginning to understand the potential risk of high doses of beta carotene and zinc, which have previously not been consumed in high doses for long periods by many people. The supplements used in this study, particularly beta carotene and zinc, are not without risk. In addition, risks may be higher among some smokers, and there is evidence from 1 epidemiologic study and 1 study of beta carotene supplementation in primates to suggest that the risks may be higher among those who drink alcohol.

High levels of zinc, a synaptic neuromodulator, can, under certain circumstances, be toxic to central neurons. Zinc damage worsens neurodegeneration in rats. This may be important in human conditions, such as transient global ischemia, that are common among elderly persons. Consistent with this, elevated zinc levels in the cerebrospinal fluid are associated with a poorer prognosis in patients following acute stroke. Elevated serum zinc is observed in patients with Alzheimer disease, and a causal role is postulated because it promotes aggregation of amyloid β-peptide. Recent data reveal that high levels of zinc, which are uptaken from the choroidal circulation, can induce apoptotic cell death of retinal pigment epithelial cells in culture. High-dose zinc supplementation also elevates glycosylated hemoglobin levels in patients with type 1 diabetes mellitus and aggravates glucose intolerance in patients with type 2 diabetes mellitus.

This raises the possibility of potential adverse health effects from high-dose supplements, a theory that needs further evaluation in long-term human studies. Until this evidence is available, a blanket recommendation for mega-dose supplements may be unwarranted.

Beyond potential risks, the very benefit of these supplements is questionable. The AREDS report continues the unfortunate trend of post hoc touting of groups exhibiting statistically significant differences among treatments. Although the α-spending function approach to group-sequential testing adjusts for repeated significance testing, it does not correct for post hoc analysis.

It is contended that for patients in categories 3 and 4, the P value for the protective effect of antioxidants plus zinc on progression to advanced AMD is .001. But considering that this particular grouping (of categories 3 and 4) is only 1 of 16 possible a priori subsets, the actual P value is \(1 - (1 - .001)^{16} = .016\), which is higher than the study’s threshold for statistical significance. The effect of combination treatment on loss of visual acuity, the only end point that truly matters to patients, which was non-significant at \(P = .02\), is even less compelling at a similarly corrected \(P = .24\).

Our anemic arsenal against AMD leads patients affected by this scourge to grasp at any straw. We owe them and their families candid advice and guidance on treatment options, even if it entails suggesting no therapy. We would recommend that the most prudent nutritional advice to offer patients with AMD is to advocate a diet rich in fruits and vegetables as opposed to any specific pharmacologic intervention, particularly a modular approach that disregards the mechanistic complexity of the disease.

Jayakrishna Ambati, MD
Lexington, Ky
Balamurali K. Ambati, MD
Durham, NC


In reply

Results from AREDS provide evidence for a clinically and statistically significant benefit of the combination of high-dose zinc and antioxidant vitamins in reducing the risk of...
developing advanced AMD in persons at high risk. This conclusion is not the result of post hoc subgroup analyses. Further review of the clinical reasons for studying various AMD categories and the statistical reasons for the analyses presented may help clarify this misperception.

The AREDS was designed as a natural history study of both AMD and cataract and as a clinical trial investigating the effects of high-dose antioxidant vitamins on slowing the development or progression of cataract and of high-dose antioxidant vitamins and zinc on slowing the development of advanced AMD. In the AREDS operations manual, written before the study began, it was estimated that persons in category 1 (few or no drusen) or category 2 (small or a few intermediate-sized drusen) would be at low risk for progression to advanced AMD during the trial. For category 1, the risk was anticipated to be so low that exposing these participants to the risks of high-dose zinc would be inappropriate. Consequently, participants in this category were not included in the AMD trial. The risk of advanced AMD for category 2 was estimated to be about 1% per year. This meant that of the planned 1000 participants in this group, we expected only 50 to develop advanced AMD during the course of the study. So with inadequate power within this group to make an assessment of the effect of treatment on the primary end point, why then did we include them in the AMD trial and expose them to the potential risks of zinc?

We recognized that if we did not include category 2 participants in the AMD trial and the trial showed that treatment was effective, this would lead to great uncertainty about what to prescribe for these patients and could have led to inappropriate treatment for many persons with small or a few intermediate-sized drusen. Note that the Physicians’ Health Study was heavily criticized for making the statistically appropriate decision not to include women physicians in their study because, at the time, there were too few women physicians older than 55 years and their risk for cardiovascular death was too low to provide useful information. Now, more than a decade later, the Women’s Health Study is trying to answer this question. Given the long duration of studies of AMD, we believed that we could not afford to be in this position at the end of our trial. One of the many goals of AREDS was to develop a severity scale for the progression of AMD similar to the scale developed for diabetic retinopathy. If an AMD scale with multiple levels could be developed, we believed that we might have enough persons in category 2 whose disease progressed sufficiently along the scale to assess possible benefits of treatment.

As it turned out, after 3 years, only 15 category 2 participants had progressed to advanced AMD. This number is far too small to assess any treatment effect on the progression to advanced AMD. We are working on developing a multilevel AMD scale. However, even without it, AREDS demonstrated large differences in the 5-year risk of progression to advanced AMD among the 4 AREDS AMD categories (1% or less for categories 1 and 2 combined vs 18% for category 3 and 43% for category 4). During the clinical trial, 316 persons progressed from category 2 to category 3 or 4. Although even for this surrogate outcome our power to find differences is limited, we found no evidence that the supplements were effective in retarding progression in persons with small or a few intermediate-sized drusen; all odds ratio estimates clustered around 1.00. Thus, including category 2 participants has provided important information about this group of patients, which constitutes a large portion of the population older than 60 years. Because they have a very low risk of progressing to advanced AMD in 5 years, and the supplements studied do not seem to slow the progression of drusen, a reasonable strategy for persons with small or a few intermediate-sized drusen may be to follow them for progression to category 3 before recommending treatment. Although including these participants in the trial has provided important clinical information, how we presented our results has created some concern about the statistical appropriateness of our recommendations.

When, in an analysis of treatment effect on progression to advanced AMD—the study’s primary outcome measure—the full cohort of participants in the AMD trial (AMD categories 2, 3, and 4) is included and the predefined design variable AMD category is a covariate in the model, the 3-degrees-of-freedom test for differential treatment effect is significant (P = .006), as are the zinc main effect (P = .009) and the treatment effect for 2 individual treatment arms, zinc alone (P = .006) and zinc plus antioxidants (P = .001). These are the same results reached by the analysis restricted to AMD category 3 and 4 participants, which are presented in the article.

Why do we adjust our statistical analyses for covariates? We do so because important covariates help explain the error (random) variability in the data and provide more precise estimates of treatment effect. The important point that is being missed by those who consider the AREDS findings the result of a post hoc subgroup analysis is this: When you have a group of participants with a negligible event rate, excluding them from the analysis (as we did in Table 4 of AREDS Report No. 8) or including them as a covariate in the model not surprisingly yields essentially the same result. We chose not to present the AMD-category–adjusted analysis for the full cohort in Table 4 of our published report because we wanted to be very clear that, although the overall results are significant, it would be inappropriate and possibly unethical to recommend supplements to a predefined group of participants who rarely have the adverse outcome of interest. There is an important distinction between this reason for exclusion and exclusion based on a non-significant treatment effect. Had we excluded the category 2 participants simply because of a nonsignificant treatment effect, we would agree with those who say the analysis was inappropriate. With the caveat that we believe that the results should not be extrapolated to the entire group studied, we include a revised Table 4 from our paper with an additional column providing the results from the analysis of all participants enrolled in the AMD trial, now with AMD category, a predefined design variable, included for covariate adjustment.

We disagree with Drs Ambati’s contention that the effect of the combination treatment on loss of visual acuity, the only end point that truly matters to patients, was nonsignificant. We believe our analysis of photographic documentation of progression to advanced AMD concurrent with a visual acuity loss of 15 letters or more is compelling evidence that the treatment has a direct impact on the disease and accompanying vision loss. The odds ratio for the combination arm on this outcome was 0.63 with a 99% confidence interval of 0.44 to 0.92 (P = .001).
Furthermore, Drs Ambati presented a P value adjustment for 16 possible a priori subsets of AMD categories. They erroneously included AMD category 1 participants who were never part of the AMD trial and were not randomized to zinc. Although there was never an intention or a rationale to study all combinations of the 3 AMD categories, had there been such intention the appropriate adjustment would be 2^3−1, or 7, possible groupings, yielding by their formula P = (1 − [1 − .001]^7) = .007. This value is in the study’s statistical significance region.

Finally, we agree with Drs Ambati, and stated in our article, that there should not be a blanket recommendation for megadose supplements. The articles they cite as evidence of potential toxicity of zinc are in vitro, case-control, short-term cohort or small randomized trials (<30 participants). The dose of zinc used in 1 of the studies of patients with diabetes mellitus is more than 8 times the dose studied in AREDS. The lack of large, well-controlled randomized trials indicates the need for further research before blanket recommendations for or against high doses of a particular nutrient can be made. We have carefully targeted the recommendation for use of AREDS-type supplements to persons at high risk of advanced AMD, with possible contraindications for smokers or other people who may have reason to avoid 1 or more of the ingredients evaluated in AREDS. As we stated in our article, AREDS participants were followed for an average of 6.3 years, and although we did not observe any serious consequences, we do not know the long-term health effects of supplementation with these high doses. Follow-up of consenting AREDS participants is continuing for at least another 5 years.

Frederick L. Ferris III, MD
Bethesda, Md
Anne S. Lindblad, PhD
Roy C. Milton, PhD
Rockville, Md

Table 4. Effect of Treatment on Risk of Progression to Advanced AMD

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Participants in AMD Categories 2, 3, and 4, Unadjusted</th>
<th>Participants in AMD Categories 3 and 4, Adjusted for AMD Category</th>
<th>Participants in AMD Categories 3 and 4, Unadjusted</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OR (99% CI)</td>
<td>P Value</td>
<td>OR (99% CI)</td>
</tr>
<tr>
<td>Antioxidants vs no antioxidants</td>
<td>0.87 (0.70-1.09)</td>
<td>.12</td>
<td>0.84 (0.67-1.06)</td>
</tr>
<tr>
<td>Zinc vs no zinc</td>
<td>0.82 (0.66-1.03)</td>
<td>.02</td>
<td>0.79 (0.63-0.99)</td>
</tr>
<tr>
<td>Antioxidants vs placebo</td>
<td>0.80 (0.59-1.09)</td>
<td>.07</td>
<td>0.75 (0.55-1.03)</td>
</tr>
<tr>
<td>Zinc vs placebo</td>
<td>0.75 (0.55-1.03)</td>
<td>.02</td>
<td>0.71 (0.52-0.98)</td>
</tr>
<tr>
<td>Antioxidants + zinc vs placebo</td>
<td>0.72 (0.52-0.98)</td>
<td>.007</td>
<td>0.67 (0.49-0.92)</td>
</tr>
</tbody>
</table>

*AMD indicates age-related macular degeneration; OR, odds ratio; and CI, confidence interval.

An Accurate Comparison of Bimatoprost’s Efficacy and Adverse Effects

I am concerned that the recent article comparing various doses of bimatoprost with both timolol maleate and a vehicle control may not fully represent all the data available to the authors. Laibovitz et al appear to only report data that are favorable to bimatoprost while neglecting other data that might produce a more balanced article.

The authors have conducted at least 2 similar studies, the first involving 60 patients (12 patients randomized to each of the treatment groups) for 1 week and the second involving 100 patients (20 patients randomized to each of the same treatment groups) who were followed for 1 month. In the study published in the July 2001 issue of the ARCHIVES, the authors do not discuss both investigations and choose to report just the second. The differences between these 2 short-term studies are minor and include dosing, duration of the study (1 week vs 1 month), and the frequency with which a technician personally administered the medications (in the 1-week study, all drops were given by a technician, and in the present study, a technician actually administered the medication only in the evening before the next study visit).

The authors (Laibovitz was the only clinical investigator in both studies, and the remaining 6 authors are long-term employees of Allergan Inc [Irvine, Calif]) did not report the data from the first, 1-week study. The 1-week study is important because subjects made more frequent scheduled visits. This schedule of follow-up visits allows for a more complete appreciation of adverse events. Additionally, compliance was insured, which may give a more realistic adverse effect profile of bimatoprost.

In the 1-week study, the adverse events of upper respiratory infections, headache, and hyperemia (Table) are far more common than in the study that the authors do report. The adverse effects from the 1-week study are more in line with the results reported in the bimatoprost package insert. This raises the following questions about the ARCHIVES article: Were the patients who experienced adverse events in the 1-week study excluded from the second, longer-term study? Do the authors believe that the differences in adverse events might have been caused by either compliance or patient selection bias?

In their discussion, the authors also do not adequately address the issues of formulation and compli-
ance. These 2 studies use nonpreserved medications rather than the commercially available preserved medication. Formulation is critical to a medication’s absorption, distribution, intraocular pressure–lowering ability, and tolerability. Likewise, the fact that medications were actually administered to the subject rather than relying on the subject’s self-administration may alter the results a clinician may expect outside this specific study situation. The authors also do not adequately inform the readers that the timolol used was a “generic” formulated by Allergan and not one formulated by Merck (Whitehouse Station, NJ).

Because of the adverse events, I am concerned about this article’s conclusion that bimatoprost is “well tolerated.” Subjects who are paid for their participation are much more likely to stay in a study than to drop out. Also, patients receiving bimatoprost are less likely to attribute symptoms of a cold or a headache to an “eye drop.” Unless clinicians are aware of the true incidence of adverse events, they may not be able to give their patients adequate informed consent or manage them appropriately. The authors should thus adequately warn readers of the symptoms and potential adverse events experienced in both studies.

Alan L. Robin, MD
Baltimore, Md

We appreciate Dr Robin’s interest in our clinical research on bimatoprost (formally AGN 192024) and agree that safety is a very important consideration when selecting a glaucoma medication.

The safety and tolerability of bimatoprost has been demonstrated by the results of the large, 1-year, pivotal studies for 0.03% Lumigan (bimatoprost ophthalmic solution; Allergan, Inc, Irvine, Calif) for 1 year. These 2 studies use nonpreserved medications rather than the commercially available preserved medication. Formulation is critical to a medication’s absorption, distribution, intraocular pressure–lowering ability, and tolerability. Likewise, the fact that medications were actually administered to the subject rather than relying on the subject’s self-administration may alter the results a clinician may expect outside this specific study situation. The authors also do not adequately inform the readers that the timolol used was a “generic” formulated by Allergan and not one formulated by Merck (Whitehouse Station, NJ).

Because of the adverse events, I am concerned about this article’s conclusion that bimatoprost is “well tolerated.” Subjects who are paid for their participation are much more likely to stay in a study than to drop out. Also, patients receiving bimatoprost are less likely to attribute symptoms of a cold or a headache to an “eye drop.” Unless clinicians are aware of the true incidence of adverse events, they may not be able to give their patients adequate informed consent or manage them appropriately. The authors should thus adequately warn readers of the symptoms and potential adverse events experienced in both studies.

In reply

The safety and tolerability of bimatoprost has been demonstrated by the results of the large, 1-year, pivotal studies for 0.03% Lumigan (bimatoprost ophthalmic solution; Allergan, Inc, Irvine, Calif). These long-term clinical data are predominantly used by the Food and Drug Administration in assessing the safety of a product and form the basis of the product labeling. These studies included 474 patients treated with bimatoprost once daily for 12 months. The 6-month data have been published;1 the 12-month data were presented at the American Academy of Ophthalmology’s 2001 annual meeting and are the source of the safety and tolerability information in the package insert.2 These results show that most treatment-related adverse events were ocular or periocular and mild in severity. The most common adverse events associated with bimatoprost were conjunctival hyperemia, eyelash growth, and ocular pruritus. After 6 months of treatment, only 11% of patients had more than mild hyperemia, and only 3% discontinued treatment because of this event. Only 7% of patients taking bimatoprost discontinued treatment because of any adverse event, which is not significantly different from the 3.3% of patients who discontinued treatment in the timolol group. There were no notable systemic findings, and the occurrences of upper respiratory infection and headaches were comparable in the bimatoprost and timolol treatment groups.

The “1-week” study that Dr Robin refers to was actually a very small (n=12 per treatment group) 5/2-day study that used twice-daily dosing (instead of the once-daily dosing stipulated in the labeling) and included treatment with a 0.1% solution—in other words, 3 times the available concentration dosed twice as often as recommended. The results of this study have been presented before3 but have little to no value in the safety assessment of bimatoprost when compared with studies of hundreds of patients receiving the drug for 1 year.

Dr Robin’s concerns regarding the protocol of the present study can be alleviated by a more careful reading of the methods section.3(p995-996) All patients were newly diagnosed (“Study Population,” fourth paragraph) and, therefore, could not have participated in the previous trial. Also, patients in the present study self-administered their own study medication at all time points except for the evenings before the 5 study visits (“Masking, Intervention, and Timing,” third paragraph). The fact that study medication was administered by a technician in the 5/2-day study only strengthens our contention that the results of that shorter study are not clinically relevant. Lastly, the timolol used in both studies was Timoptic (Merck, Whitehouse Station, NJ), not a generic.

Amanda M. VanDenburgh, MS, MBA
Irvine, Calif


### Number (Percentage) of Subjects With Adverse Effects During a 1-Week Study*

<table>
<thead>
<tr>
<th>Adverse Effect</th>
<th>Bimatoprost 0.01% BID (n = 12)</th>
<th>Bimatoprost 0.03% BID (n = 12)</th>
<th>Bimatoprost 0.03% Package Insert (n = 474)</th>
<th>Bimatoprost 0.1% BID (n = 12)</th>
<th>Timolol Maleate BID (n = 12)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Upper respiratory infection</td>
<td>2 (17)</td>
<td>1 (8)</td>
<td>47 (10)</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Headache</td>
<td>3 (25)</td>
<td>1 (8)</td>
<td>5–24 (1-5)</td>
<td>4 (33)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Hyperemia</td>
<td>1 (8)</td>
<td>5 (42)</td>
<td>213 (45)</td>
<td>8 (67)</td>
<td>2 (17)</td>
</tr>
</tbody>
</table>

*Data are from the Center for Drug Evaluation and Research unless otherwise indicated. BID indicates twice daily.