Visual Acuity Outcomes Among Sham vs No-Treatment Controls From Randomized Trials

Barbara S. Hawkins, PhD; Neil M. Bressler, MD; Sandra M. Reynolds, MA

Objective: To compare 2-year visual acuity outcomes between similar participants assigned to sham and no-treatment control arms in randomized clinical trials.

Methods: We retrospectively matched sham controls from 2 randomized trials to no-treatment controls (no sham or placebo) from 3 trials on 8 baseline prognostic criteria (full matches) or on 4 to 7 criteria (partial matches). Outcomes were compared using data from those who had 2-year visual acuity measurements and also using the last observation carried forward method to impute missing 2-year measurements.

Results: A full match to a no-treatment control was identified for 72 of 321 sham controls (22%); a partial match was identified for another 93 sham controls (29%). Among the fully matched pairs, no important difference in 2-year visual acuity outcomes was observed. However, 2-year outcomes differed somewhat between sham and no-treatment controls within the partially matched pairs.

Conclusions: Findings from fully matched pairs suggest that sham treatment to mask participants in clinical trials may be unnecessary when visual acuity is the outcome of interest. However, findings from the partially matched pairs do not fully support this conclusion. This analysis challenges the necessity for sham (placebo) controls in randomized clinical trials in ophthalmology when visual acuity is the primary outcome of interest.

Arch Ophthalmol. 2009;127(6):725-731

As Kapchuk observed, “In the current RCT [randomized controlled trial] era, a legitimate therapy must demonstrate an effect greater than a decoy disguised as a real intervention.” In RCTs of pharmacologic agents administered either orally or topically, placebos (one type of sham or decoy control) composed of physiologically inert or inactive material and matched to the test agent with respect to appearance, taste, or other characteristics have become an accepted control. Reasons to use sham controls include facilitation of outcome masking, promotion of similar compliance in each arm of the trial, control for natural history of the medical condition, control for nontrial interventions, and regulatory requirements. Reasons to avoid sham controls include inaccurate estimation of effect sizes and adverse event rates, interference with nonspecific aspects of treatment, and failure to inform clinical decisions.

Ingestion or topical application of an inactive or inert placebo poses little or no risk to the participant in a clinical trial. However, there are costs associated with developing, producing, and distributing the matching placebo and additional costs associated with monitoring adherence to the schedule for ingestion or application of the placebo agent, similar to those for the test agent. In some situations, such as those requiring ingestion or topical application several times per day or ingestion of a large number of placebo tablets or capsules to match the regimen for the test agent, the participant assigned to the placebo arm may be inconvenienced with no chance of benefit.

For editorial comment see page 801

Although sham laser photocoagulation, sham injections without needles, sham radiotherapy, intravenous injection of a placebo solution such as dextrose, or intravitreous injection of saline may have low associated risk, intraocular injection of a vehicle without an active agent or implantation of a device designed for local drug delivery without the drug may have a risk that a trial participant would not be willing to accept unless convinced that there was potential personal benefit. Sham controls of some types have been questioned on both scientific and ethical grounds. Sham procedures not only may entail risk but also may be inconvenient to the participant in the sham...
To our knowledge, the question of preference for sham vs no treatment has not been posed to potential or actual participants in clinical trials in ophthalmology.

These issues have become increasingly important in ophthalmology with the development of new interventions and delivery systems for ophthalmologic conditions. Apart from a single crossover trial in dry eye syndrome, no clinical trial in any ophthalmologic condition has compared sham (placebo) treatment with no treatment among participants randomly assigned to the control arm, although such comparisons have been made in clinical trials in other medical specialties. From their updated systematic review of 156 clinical trials in which participants were assigned randomly to sham (placebo) or no-treatment control arms and in which there was masked assessment of objective outcomes, Hróbjartsson and Gøtzsche were unable to detect a statistically significant overall effect of placebo interventions in trials with binary outcomes, whether reported by patients or observers, or in trials with continuous outcomes reported by observers. A statistically significant moderate difference between placebo and no-treatment groups was observed overall for trials with continuous outcomes reported by patients, and for trials involving patient-reported pain and phobia.

The findings from this systematic review prompted us to ask whether sham controls were essential in RCTs with outcomes based on visual acuity. In the absence of RCTs with visual acuity outcomes in which participants in the control arm have been assigned randomly between a placebo or sham treatment and no treatment, we sought databases from clinical trials in patients with similar retinal conditions, some of which used a sham control arm and others a no-treatment control arm, so that we could compare visual acuity outcomes. We evaluated the influence of sham controls on 2-year visual acuity outcomes, including completion of 2-year visual acuity examinations, using matched pairs of sham and no-treatment controls. Our goal is to stimulate research into the use of sham controls so that scientifically and ethically sound recommendations can be formulated to inform regulatory requirements for future clinical trials in ophthalmology.

### METHODS

#### SELECTION OF PATIENTS

We obtained from the sponsors or coordinating centers data for patients in the control arms of 5 completed randomized trials of treatments for subfoveal choroidal neovascularization (CNV) secondary to age-related macular degeneration (AMD). We obtained permission to use the control arm data for this analysis from the responsible committees or investigators in each trial. No personal identifying information was requested or provided. In 2 of the trials, a sham control (intravenous injection of placebo followed by low-energy laser) was used with the goal of masking the participant, the treating ophthalmologist, and the visual acuity examiner.

In 3 of the trials, a no-treatment (observation) control was used so that the participant and the ophthalmologist were unmasked. In 1 of the trials with no-treatment controls, the visual acuity examiner at the 2-year examination was masked. In all of the 5 trials, the neovascular lesion had to be in an eye with AMD and had to be located under the geometric center of the foveal avascular zone. Table 1 summarizes other characteristics of the 5 trials and eligibility criteria pertinent to this investigation.

We attempted to match each sham-treated participant to 1 or more participants in the no-treatment arms using baseline characteristics reported to be predictive of 2-year changes in visual acuity: gender, age (within 2 years), visual acuity of the study eye (within 7 letters, ie, 0.14 logMAR), size of the neovascular lesion (same category: ≥2, >2 but ≤6, and >6 disc areas), classic CNV in the study eye (present or absent), CNV in the fellow (nonstudy) eye (present or absent), blood as part of the neovascular lesion (present or absent), and prior thermal laser treatment in the study eye (yes or no). Initially, the matching criterion selected for visual acuity was more stringent, ie, ≥2 letters. However, few pairs of controls could be matched on visual acuity and other characteristics using this stricter criterion.

After the best possible match on all of the 8 criteria had been identified for as many sham-treated participants as possible, the matching criteria except those regarding baseline age and visual acuity of the study eye were removed one at a time and the best matches on 7 criteria were identified from the remaining participants. After the best matches on 7 criteria had been identified, pairs of matching criteria were removed and the best matches on 6 criteria were identified from the remaining participants. This process was repeated until all of the possible matches were identified by applying as many criteria as possible in addition to baseline age and visual acuity of the study eye.

Methods similar to those described earlier were used in a secondary analysis of the effect of masked visual acuity examiners. Both trials from which control arm patients were selected for this analysis included no-treatment controls; patients were unmasked in both trials. In one of the trials, 2-year visual acuity measurements were made by masked examiners; in the other trial, no effort was made to mask the visual acuity examiners. The baseline lesion size categories used to match masked and unmasked control arm patients were modified to conform to more precise classification of size in these 2 trials.

The study was approved by one of the institutional review boards (IRBX) of the Johns Hopkins Medical Institutions.

#### DATA ANALYSIS

We compared distributions and means of 2-year changes in visual acuity (continuous outcome) between sham and no-treatment controls. We also compared the groups with respect to dichotomous 2-year outcomes evaluated in 1 or more of these 5 trials as indicating failure of the test treatment, ie, loss of 3 or more and loss of 6 or more lines of visual acuity from baseline levels. When a sham control matched more than 1 no-treatment control, data from the no-treatment control who matched the sham control most closely were used in analyses. Findings in each group of matched controls have been analyzed both for study eyes examined at 2 years (available cases) and for all of the study eyes with the last observation carried forward (LOCF) method to impute missing 2-year visual acuity measurements. For 1 sham control with no earlier fol-
low-up measurement of visual acuity, the 3-year measurement was used for the LOCF analyses.

Distributions have been displayed using box-and-whisker plots and summary statistics. Wilcoxon matched pairs signed rank sum tests were used to assess the similarity of distributions, and paired t tests were used to assess the similarity of mean values. McNemar tests were used to compare dichotomous (binary) outcomes among matched pairs of controls. Linear regression models were used to evaluate the effect of discrepancies in matching criteria between partially matched pairs of controls on 2-year changes in visual acuity.

We used SAS version 8.2 statistical software (SAS Institute, Cary, North Carolina) for all of the analyses. We did not adjust probabilities to account for multiple comparisons; probability values of .05 or less were deemed indicators of noteworthy differences between the sham and no-treatment controls.

## SUCCESS OF MATCHING

Of the 321 sham controls, matches on all of the 8 criteria (full matches) among no-treatment controls were found for 72. Of the remaining 249 sham controls, matches on 4 to 7 criteria (partial matches) were found for 93. Of the 93 pairs, 46 matched on 7 criteria, 34 matched on 6 criteria, 12 matched on 5 criteria, and 1 matched on 4 criteria. No additional matches on 3 criteria (baseline age and visual acuity plus any other criterion) or on baseline age and visual acuity alone were possible. Of the 136 unmatched sham controls, 92 had no classic CNV in the subfoveal lesion and were primarily from the Verteoporfin in Photodynamic Therapy Trial for which classic CNV as part of the subfoveal lesion was not required (Table 1). Of the remaining 44 unmatched sham controls, 17 either did not have a subfoveal lesion or did not have AMD based on assessment of photographs at the central reading center. The study eye visual acuity of all other unmatched sham controls was better than that of all remaining no-treatment controls.

Of the 72 sham controls in the fully matched group and 93 sham controls in the partially matched group, 70 and 88, respectively, were from the Treatment of Age-Related Macular Degeneration With Photodynamic Therapy Investigation. The sources (clinical trials) of no-treatment controls in the 2 matched groups are shown in eTable 1 (http://www.archophthalmol.com).

Despite matching, the distributions of study eye visual acuity scores at baseline (Figure 1) differ between the sham and no-treatment controls within both the fully matched and partially matched groups (Wilcoxon and t test P < .001, both comparisons). The mean difference in visual acuity score is 2.2 letters among the fully matched pairs and 2.6 letters among the partially matched pairs (Figure 1B). In both groups, no-treatment controls have somewhat poorer visual acuity. The paired mean difference in age is 0.1 year in both groups.

The most frequently discrepant criterion between sham and no-treatment controls in the partially matched group, affecting 49 of the 93 partially matched pairs of controls, was the size of the subfoveal neovascular lesion despite the broad size categories. The number of partially matched pairs discrepant with respect to other criteria ranged from 8 (presence or absence of classic CNV) to 31 (gender).

## COMPLETENESS OF 2-YEAR EXAMINATIONS

Among the fully matched group, 8 of the 72 sham controls (11%) missed the 2-year examination compared with 10 of the no-treatment controls (14%). Among the partially matched group, 14 of the 93 sham controls (15%)...
COMPARISON OF VISUAL ACUITY OUTCOMES

Distributions of changes in visual acuity from baseline to 2 years are shown in Figure 2A for patients who completed 2-year visual acuity examinations. The distributions of the differences in 2-year changes in visual acuity between matched pairs of controls are shown in Figure 2B. The LOCF distributions (data not shown) were nearly identical to those shown for available cases.

Comparisons of mean changes in visual acuity along with dichotomous visual acuity outcomes are summarized in Table 2 for fully matched pairs of controls. The absolute mean difference was 0.16 line, ie, less than 1 letter, when patients who completed a 2-year examination were analyzed. The LOCF method yielded an absolute mean difference of 0.10 line. No 2-year visual acuity outcome differed substantially between sham and no-treatment controls in the fully matched group regardless of the method of analysis used or the source of no-treatment controls (all P > .15).

For the partially matched group (Table 3), the mean 2-year change was 3.5 lines lost among sham controls and 5.0 lines lost among no-treatment controls who completed the 2-year examination (paired t test P = .02); the absolute mean difference was 1.5 lines (7.5 letters). When the last observation was carried forward to 2 years, the mean changes and absolute mean difference were similar to those of eyes examined at 2 years. Outcomes among

Table 2. Comparison of 2-Year Visual Acuity Outcomes Between Fully Matched Pairs of Sham and No-Treatment Controls

<table>
<thead>
<tr>
<th>2-y Visual Acuity Outcome</th>
<th>Sham Controls (n=72)</th>
<th>No-Treatment Controls (n=72)</th>
<th>Mean Difference, Sham Control−No-Treatment Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eyes with ≥ 3 lines lost, No. (%)</td>
<td>Observed: 4.0 (20)</td>
<td>4.2 (21)</td>
<td>−0.16 (0.8)</td>
</tr>
<tr>
<td>LOCF: 4.2 (21)</td>
<td>4.1 (21)</td>
<td>0.10 (0.5)</td>
<td></td>
</tr>
<tr>
<td>Eyes with ≥ 6 lines lost, No. (%)</td>
<td>Observed: 38 (68)</td>
<td>33 (59)</td>
<td>9%</td>
</tr>
<tr>
<td>LOCF: 50 (69)</td>
<td>41 (57)</td>
<td>12%</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviation: LOCF, last observation carried forward.

a Observed indicates study eyes with 2-year visual acuity measurements (available cases, 56 pairs). The LOCF method imputes 2-year visual acuity for eyes not examined at 2 years (72 pairs).

and 11 of the no-treatment controls (12%) missed the 2-year examination. Thus, there was no evidence that masking participants and clinical personnel via sham treatment played a major role in completeness of follow-up for visual acuity outcomes.

Of the 11 no-treatment controls in the partially matched group whose visual acuity was not measured at 2 years, one had died 10 days after enrollment in the Submacular Surgery Trials (SST) and another had no visual acuity measurement during follow-up. Neither of these
partially matched pairs were worse for no-treatment controls than for sham controls (Table 3), which is in contrast to the fully matched pairs where outcomes were similar or slightly better among no-treatment controls (Table 2). No pairwise discrepancy resulting from relaxing matching criteria was found to influence differences between types of partially matched controls when evaluated using linear regression models.

The same outcomes were analyzed separately for partially matched pairs of sham and no-treatment controls by the source of no-treatment controls. Most of the differences in 2-year visual acuity outcomes were small when no-treatment controls were from the SST (eTable 2) (all \( P > .30 \)). However, when no-treatment controls were from the Macular Photocoagulation Study (MPS), both mean change in visual acuity and loss of 3 or more lines of visual acuity differed between the 2 types of controls regardless of the analytic strategy (all \( P < .05 \)). The differences were less extreme for 2-year loss of 6 or more lines (\( P = .19 \) and \( .09 \) for available cases and LOCF analyses, respectively). One possible explanation for this difference in findings in these 2 subgroups is that almost all of the SST patients had visual acuity measured at the 2-year examination by a masked examiner, whereas no attempt was made to mask MPS visual acuity examiners.

**EFFECT OF MASKED VISUAL ACUITY EXAMINER ON OUTCOMES AMONG UNMASKED PATIENTS**

Sixty-eight pairs of no-treatment controls from the SST who had masked 2-year visual acuity examinations were matched to no-treatment controls from the MPS. Data from fully matched (n=4) and partially matched (n=64) pairs of no-treatment controls were pooled to analyze the effect of masking the visual acuity examiner on 2-year change in visual acuity. No-treatment controls from the MPS lost a mean of 1.3 lines (6.5 letters) more than matched controls from the SST in the 2-year interval following enrollment and treatment when only available cases (55 pairs) were considered and lost 1.6 lines (8 letters) more when using the LOCF approach (68 pairs). These differences are almost as large as those observed when no-treatment controls from the MPS were matched to sham controls (eTable 2).

In this retrospective comparison of 2-year visual acuity outcomes in sham vs no-treatment controls who participated in randomized controlled trials of treatments for subfoveal CNV secondary to AMD, the estimated difference (placebo effect) was small or trivial among fully matched pairs regardless of which of the visual acuity outcomes was considered. The estimated difference was somewhat larger among the partially matched pairs, particularly when analyses used the LOCF method to impute missed 2-year visual acuity measurements. Our retrospective analysis of visual acuity outcomes among the full matches supports the conclusions of Hrobjartsson and Gøtzsche,21,22 who found in their review of clinical trials in various medical conditions “no evidence that placebo interventions in general have clinically important effects.”21 However, the findings for partial matches are inconsistent with those from full matches. We are aware of only 1 other comparison of sham and no-treatment controls from a clinical trial of treatment of subfoveal CNV secondary to AMD.30 Some centers participating in that clinical trial used sham radiotherapy controls, and others used no-treatment controls; the investigators reported that the treatment effect did not differ by type of control.30

Our investigation was limited to a single ophthalmologic diagnosis among control arm participants from 5 of the many RCTs now completed for this condition. Visual acuity, a quasi-objective measurement, may be more robust than other outcomes. Nevertheless, cooperation by the patient and patience on the part of the examiner are required to obtain accurate, reproducible measurements, especially for patients with poor vision. Although personnel at the same reading center described the characteristics of subfoveal neovascular lesions, there were differences in the way data available for this analysis had been recorded owing to differences in the goals and treatment protocols for the 5 trials. Thus, some matching criteria were less precise than others.

Although matched sham controls were primarily from a single clinical trial, no-treatment controls were from 2 different studies. Thus, we had an opportunity to examine matched pairs by source of no-treatment controls. No difference by source of no-treatment controls was found for fully matched pairs. However, for partially matched pairs, differences in outcomes between matched pairs tended to be larger when no-treatment control matches were from the MPS than when they were from the SST. As noted earlier and in Table 1, the MPS was concluded before the other clinical trials were initiated. Neither MPS patients nor MPS

---

**Table 3. Comparison of 2-Year Visual Acuity Outcomes Between Partially Matched Pairs of Sham and No-Treatment Controls**

<table>
<thead>
<tr>
<th>Mean Difference, Sham Control− No-Treatment Control</th>
<th>Eyes with ( \geq 6 ) lines lost, No. (%)</th>
<th>LOCF</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-y Visual Acuity Outcome</td>
<td>Observed</td>
<td>22 (31)</td>
</tr>
<tr>
<td></td>
<td>LOCF</td>
<td>31 (43)</td>
</tr>
</tbody>
</table>
visual acuity examiners were masked to the random assignment during follow-up, possibly accounting for these differences. A secondary analysis of 68 matched pairs of controls suggested that masking of the visual acuity examiner may have been responsible for part of the difference in 2-year outcomes observed in the partially matched group of sham and no-treatment controls. However, it is not possible to conclude that masked visual acuity examiners rather than temporal changes in interpretation of fluorescein angiograms or other factors are the primary explanation for the differences observed.

The 5 clinical trials from which data were obtained for this analysis were of high quality. In all of the 5 trials, visual acuity examiners followed a standard protocol to obtain the best refractive correction and to measure visual acuity at baseline and follow-up examinations. All of the 5 trials incorporated training and regular monitoring of visual acuity examiners. A high proportion of enrollees completed the 2-year visual acuity examination, so 2-year outcomes based on analysis of available cases were similar to those from analysis using the LOCF method. The same photograph reading center participated in all of the 5 trials; thus, similar methods were used by graders to record the characteristics of the neovascular lesions.

This retrospective analysis cannot replace randomized comparisons of sham and no-treatment controls in clinical trials of treatment of subfoveal CNV in AMD. Because treatments that usually delay, prevent, or reverse visual acuity loss are currently available, it is unlikely that there will be an opportunity to make a randomized comparison in future trials of treatments for patients with this condition. However, we encourage sponsors and designers of clinical trials of new therapeutic and preventive interventions for ophthalmologic conditions for which no effective treatment is available to consider a design in which patients assigned to the control arm are randomized between sham and no treatment regardless of whether visual acuity or some other outcome is assessed. A modification of such a design could be used in clinical trials in which a new treatment is to be evaluated as an adjunct to a standard treatment; in such cases, 2 types of sham controls have been used, perhaps unnecessarily. Multiple trials with an internally randomized control arm would yield a body of evidence that should help to delineate the circumstances that require sham controls to minimize bias in estimating treatment effects along with the risk, cost, deception, and inconvenience they entail and those in which no-treatment controls are acceptable. Such a body of evidence would be instructive to researchers and to regulatory agencies that approve clinical trial protocols and new interventions. Furthermore, the goals of individual clinical trials must be considered. As Vickers and de Craen noted in their review, trials with no-treatment controls “may provide information that is more immediately applicable to clinical practice.”

Until sufficient evidence is available to inform decisions regarding selection of the control “treatment,” we encourage designers of clinical trials in ophthalmology to resist the presumption of superior scientific advantages of sham controls. All scientific, ethical, logistic, and financial aspects of masking participants should be considered before deciding whether a sham control or a no-treatment control, together with other protections against bias such as masked outcome assessment, is the best method to achieve trial objectives.

Submitted for Publication: October 7, 2008; final revision received December 23, 2008; accepted December 29, 2008.

Correspondence: Barbara S. Hawkins, PhD, Wilmer Eye Institute, 550 N Broadway, Room 930, Baltimore, MD 21205-2010 (bhawkins@jhmi.edu).

Author Contributions: Dr Hawkins had full access to the databases used for the analyses reported herein.

Financial Disclosure: None reported.

Funding/Support: This study was supported in part by the Karl P. Hagen Professorship in Ophthalmology (Dr Hawkins), the James P. Gill Professorship in Ophthalmology (Dr Bressler), the Retina Division Research Fund of the Wilmer Eye Institute, and an unrestricted grant to the Wilmer Eye Institute from Research to Prevent Blindness, New York, New York.

Trial Registration: The Submacular Surgery Trials are registered at clinicaltrials.gov under identifier NCT00000150, and the Macular Photocoagulation Study is registered at clinicaltrials.gov under identifier NCT00000158.

Previous Presentations: This work was presented in part at the 28th Annual Meeting of the Society for Clinical Trials; May 21, 2007; Montreal, Quebec, Canada; and the 31st Annual Meeting of the Macula Society; March 29, 2008; Palm Beach, Florida.


Additional Contributions: The sham control databases were provided by QLT, Inc, Vancouver, British Columbia, Canada. Permission to use the no-treatment control databases was provided by the Macular Photocoagulation Study Legacy Committee and the Submacular Surgery Trials Archives Committee. The Macular Photocoagulation Study and the Submacular Surgery Trials were sponsored by the National Eye Institute, National Institutes of Health, US Department of Health and Human Services.

REFERENCES


From the Archives of the Archives

140 Years Ago . . .

About fifteen years ago the genius of Prof. Virchow predicted that, with the ophthalmoscope, embolism in the retinal artery might be directly seen in the living body. This suggestion was a fruit of his brilliant discoveries of the varied series of morbid changes, resulting from the obstruction of blood vessels by thrombosis and embolism. Four years later, Prof. V. Graefe observed the first case in which almost instantaneous blindness was caused by obstruction of the central retinal artery, in a patient suffering from indocarditis [sic].