



Published in final edited form as:

*Am J Ophthalmol.* 2014 July ; 158(1): 96–104.e2. doi:10.1016/j.ajo.2014.04.003.

## Visual Field Improvement in the Collaborative Initial Glaucoma Treatment Study

David C. Musch<sup>1,2</sup>, Brenda W. Gillespie<sup>3</sup>, Paul F. Palmberg<sup>4</sup>, George Spaeth<sup>5</sup>, Leslie M. Niziol<sup>1</sup>, and Paul R. Lichter<sup>1</sup>

<sup>1</sup>Department of Ophthalmology and Visual Sciences, Medical School, University of Michigan, Ann Arbor, MI

<sup>2</sup>Department of Epidemiology, School of Public Health, University of Michigan, Ann Arbor, MI

<sup>3</sup>Department of Biostatistics, School of Public Health, University of Michigan, Ann Arbor, MI

<sup>4</sup>Bascom Palmer Eye Institute, University of Miami, Miami, FL

<sup>5</sup>Wills Eye Hospital, Philadelphia, PA

### Abstract

**Purpose**—To evaluate critically visual field (VF) improvement in participants in the Collaborative Initial Glaucoma Treatment Study (CIGTS).

**Design**—Prospective, comparative case series from a randomized clinical trial comparing trabeculectomy and topical medications in treating open-angle glaucoma (OAG).

**Methods**—607 subjects with newly-diagnosed OAG were identified for study. Baseline and follow-up VF tests were obtained and mean deviation (MD) change from baseline over follow-up was analyzed. Clinically substantial change (loss or improvement) was defined as change from baseline of  $\geq 3$  decibels in MD. Baseline factors were inspected to determine their association with VF improvement in repeated measures regression models.

**Results**—The percentage of participants showing substantial VF improvement over time was similar to that showing VF loss through five years after initial treatment, after which VF loss became more frequent. Measures of better intraocular pressure (IOP) control during treatment were significantly predictive of VF improvement, including a lower mean IOP, a lower minimum IOP, and lower sustained levels of IOP over follow-up. Other predictive factors included female sex [odds ratio (OR)=1.73], visits one year prior to cataract extraction (OR=0.11), and an

© 2014 Elsevier Inc. All rights reserved.

Corresponding Author: David C. Musch, Ph.D., M.P.H., University of Michigan – Kellogg Eye Center, 1000 Wall Street, Ann Arbor, MI 48105, Telephone: 734-763-8175, Fax: 734-232-8181, dmusch@med.umich.edu.

**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.

#### Author Contributions

Design and conduct of the study (DCM, BWG, PFP, PRL); Collection, management, analysis, and interpretation of the data (DCM, BWG, PFP, GS, LMN, PRL); preparation, review, or approval of the manuscript (DCM, BWG, PFP, GS, LMN, PRL).

Supplemental Material available at [AJO.com](http://AJO.com)

interaction between treatment and baseline MD wherein surgically treated subjects with worse baseline VF loss were more likely to show VF improvement.

**Conclusions**—In the CIGTS, substantial VF loss and improvement were comparable through five years of follow-up, after which VF loss became more frequent. Predictive factors for VF improvement included several indicators of better IOP control, which supports the postulate that VF improvement was real.

---

## Introduction

Visual field (VF) loss is a hallmark sign of glaucoma, and its assessment is a standard part of the ophthalmic examination of patients being evaluated or followed for this condition. Measurement of VF loss is performed using some form of VF testing, which in current practice is commonly an automated test. The test involves presenting a sequence of light stimuli to a patient and judging the patient's ability to detect the stimuli. Its reliability is estimated by recording false positive and false negative responses as well as fixation losses which, if substantial, require re-testing. Within-patient variability in the VF test is a well-known phenomenon,<sup>1,2</sup> and its study has resulted in the recommendation that a finding on one test be confirmed by subsequent repeated testing.<sup>3</sup> For example, in three major trials of glaucoma and ocular hypertension treatment,<sup>4-6</sup> three consecutive VF tests were required to confirm a defined change in VF.

Most studies involving treatment of glaucoma have evaluated progression in VF loss to measure treatment efficacy. Our previous investigations using the Collaborative Initial Glaucoma Treatment Study (CIGTS) data have provided information on factors predictive of progressive VF loss.<sup>7,8</sup> The possibility that VF improvement might occur has received sparse attention. In 1985, Spaeth<sup>9</sup> presented evidence that lowering intraocular pressure (IOP) in patients with glaucoma can result in improvement of VFs. He suggested the possibility that retinal ganglion cells damaged by glaucoma may, with effective treatment, recover their function; thus, not only VF loss but also VF improvement should be considered as real phenomena. If so, VF improvement would not necessarily be an artifact of inherent noise in test taking, which has been the prevailing thought.

The purpose of this study was to evaluate the distribution of VF change over 9 years of follow-up in CIGTS, with a particular focus on those VF tests that demonstrated improvement from a carefully measured baseline. We sought to assess whether improvement was associated with measures indicative of good IOP control, based on the premise that factors previously identified as predictive of VF loss should also be found to associate (in the opposite direction) with VF improvement, if improvement is real.

## Methods

The CIGTS was a multicenter, randomized clinical trial in which 607 participants with newly diagnosed glaucoma were assigned to initial treatment with topical medications or trabeculectomy. For bilaterally eligible participants, prior to randomization a study eye was selected by the enrolling ophthalmologist to receive the randomized treatment. The VF findings evaluated herein relate to the study eye of each patient. Patients were enrolled

between 1993 and 1997 and follow-up continued through 2004. Informed consent to participate was obtained from all patients and IRB approval was obtained at all participating centers. The research was HIPPA compliant and adhered to the tenets of the Declaration of Helsinki.

Participants were required to have had at least one VF test prior to being screened for eligibility, and had two comprehensive baseline examinations that included VF testing. The Humphrey 24-2 full threshold VF test was administered by personnel who were certified and followed a well-defined protocol for the test. If the two baseline VF tests varied by more than three units in the scoring or if one was not reliable, a third VF test was obtained. The average of the two tests, or the median of three, was used to characterize the participant's baseline VF status. Using the same VF test protocol, VF tests were performed at follow-up visits conducted at 3 and 6 months after treatment initiation, and bi-annually thereafter. Further details on the study's protocol and participants have been previously reported.<sup>5</sup>

Substantial progression in the extent of VF loss during the CIGTS was a trigger for instituting more aggressive treatment. The amount of change required to warrant further treatment was determined prior to the study's initiation by convening a panel of glaucoma experts who were presented with side-by-side VFs that showed various increments of VF progression, as measured by the CIGTS VF score. This score has been described previously.<sup>2</sup> In brief, its calculation relied upon the probabilities in the 52 locations of the total deviation probability plot, with weighting assigned based on the number of affected adjacent points. The result is a global measure of VF loss on a scale from 0 to 20, with higher scores indicative of more loss. The expert panel determined by consensus that an increment of 3 units in the CIGTS VF score represented a clinically substantial change. In subsequent analyses, a standard output of VF testing, the mean deviation (MD) was found to be highly correlated with the CIGTS VF score, and the MD was also a more sensitive measure of substantial VF defects than the CIGTS VF score. Thus, for this study, we employed a 3 dB change in MD as the minimum incremental change of clinical relevance.

### Statistical methods

The percentages of CIGTS participants who demonstrated at least a 3 dB improvement (gain) in their MD from baseline were plotted by follow-up visit, along with those who demonstrated at least a 3 dB loss for comparison. To address sustained VF gain or loss, we identified visits with gain or loss that were validated by a subsequent gain or loss at their next visit 6 months later. Those without a subsequent six month value were not included in either the numerator or denominator at a given study visit point. Standard errors of the binomial proportion estimates ( $p$ ) were calculated as  $\sqrt{p*(1-p)/n}$ , where  $p$  is the proportion of subjects showing gain/loss. For those with sustained gain, we also plotted their MD values over time in spaghetti plots to check consistency of the gains at subsequent time points (see supplementary Figure 1; supplemental material at [AJO.com](http://AJO.com)).

To investigate whether consecutive measurements showing gain over 5 years of follow-up are consistent with chance occurrence, we simulated the number expected to have gain under the assumption of no real VF change over time, accounting for correlation between

visits over time, and compared it with the observed number with gain. Details of this calculation are given in a footnote to the relevant table.

We used repeated measures logistic regression to evaluate factors associated with VF improvement, where the dichotomous outcome was at least a 3 dB gain in MD from baseline at each time point. For example, if a subject had a baseline MD of  $-5.0$  dB and had a MD at years 1,2,3, and 4 of  $-2.0$ ,  $-1.5$ ,  $-1.0$ , and  $-4.0$  dB, respectively, the repeated measures outcome for those years would be 1, 1, 1, and 0, where 0 = no improvement and 1= at least a 3 dB improvement. A robust variance estimate<sup>10</sup> was used to account for repeated MD measures in an eye over time. This approach permitted us to evaluate the association of baseline variables with the VF improvement outcome as well as time dependent (during follow-up) associations of IOP control with VF improvement. Baseline covariates tested in the model included MD, visual acuity (VA), IOP, corneal thickness, treatment, age, sex, race, marital status, diabetes, hypertension, other vascular/cardiac disease, type of open-angle glaucoma, family history of glaucoma, disc hemorrhage, pupillary response, iris color, smoking status, alcohol consumption, and center. Time-dependent covariates investigated included visit number and IOP values (based only on IOP measurements prior to each time point): mean; standard deviation (SD); minimum; maximum; and range of IOP measures; percent of visits with IOP less than 16, 18, 20, or 22mmHg; and indicators for whether all visits had IOP less than 16, 18, 20, or 22mmHg. Cataract was accounted for in the model as a time-dependent variable indicating a visit within 1 year before cataract surgery to capture the VF decrement and subsequent lower chance of improvement during this period. Our model selection strategy included a best subset selection method ignoring the correlation structure, followed by testing the best of the identified models in the repeated measures setting. This strategy included testing single variable models. Interactions of all baseline covariates with treatment and time were investigated, whether or not the main effects were significant. Models were run on follow-up data from years 2–9. Starting at year 2 insured at least two years of summary IOP data to inform those covariates. Cox regression based on time to sustained VF gain was used to identify baseline and time-dependent IOP effects (detailed above) associated with time to first sustained VF gain. Proportional hazards were tested with covariate by time interactions. SAS version 9.3 statistical software (SAS, Cary, NC) was used.

## Results

A total of 607 subjects were enrolled in the CIGTS. A description of the sample is given in Table 1. Briefly, subjects were on average 58 years old at study entry, 55% male, 56% white and followed for a mean of 7.2 years. Clinical characteristics at baseline included an average MD of  $-5.4$  dB, IOP of 27.5 mmHg, and VA of 85.7 ETDRS letters (85 letters equates to a 20/20 Snellen VA).

The percentages of CIGTS participants' study eyes that demonstrated substantial change from baseline – either gain or loss – in study follow-up through five years was very similar (Figure 1). For example, at 1, 3, and 5 years after treatment initiation, the percentage showing loss (6.6%, 10.9%, and 14.5%, respectively) and improvement (7.5%, 12.7%, and 13.9%, respectively) differed minimally and were not statistically significant ( $p>0.20$ ) and

the increasing trends in both loss and improvement were similar. Between 5 and 8 years, the percentage with VF loss continued to increase whereas VF improvement remained at the 5 year level, resulting at 7 years after treatment initiation in a higher percentage of patients showing VF loss (19.6%) than improvement (13.5%). The percentage of VFs showing a 3 dB gain was significantly different from the percentage showing comparable loss at years 7 and 8 (test of equality of binomial proportions; p-values=0.0053 and 0.0088, respectively). Although the percentages of subjects showing a sustained gain or loss are smaller, a similar pattern is present with 3.9%, 7.0%, and 7.2% % showing sustained gain and 2.9%, 7.3%, and 11.0% showing sustained loss at 1, 3, and 5 years after treatment initiation, respectively.

The tests of equal probability of VF improvement and loss at each time point cannot validate that either loss or gain is real. Thus, we investigated three potential alternative explanations for the observed improvement. These included learning effects, variability in MD measures, and chance occurrence in long-term follow-up. *Learning effects* were evaluated by revisiting previously reported conclusions based on MD values from the two baseline CIGTS visits.<sup>2</sup> 82% of these visits were conducted within 20 days of each other, and all took place within 42 days of each other. These CIGTS results revealed a small, but significant, learning effect (the second VF was an average of 0.57 dB better than the first) for VFs performed within 20 days of each other, but no significant learning effect when the second VF was conducted more than 20 days after the first. As the first follow-up visits took place 3 months after treatment initiation, all follow-up visits took place >20 days from baseline.

*Variability* was estimated by calculating the percent of patients with 3dB loss or gain at the second baseline VF compared to the first. Out of 605 where the order of fields could be determined, 25 (4.1%) showed loss. Of the 110 fields that were obtained more than 20 days apart (to avoid learning effects), 5 (4.6%) showed gain. Thus, we estimate that approximately 4% of supposed losses or gains may be attributable to variability in VF measurement in the absence of any real change in VFs.

*Chance occurrence* of consistent gain or loss in long-term follow-up was first evaluated by inspecting a spaghetti plot of VF MD values for subjects with MD improvement from baseline of at least 3 dB verified by improvement at the next clinical visit (6 months later) (n=107). This plot shows the gains were for the most part maintained through follow-up (supplemental material at [AJO.com](http://AJO.com)). We also looked at the distribution of consecutive follow-up visits showing gain or loss over the first 5 years compared to what would be expected by chance alone, when assuming no VF change over time and correlation between visits. These values are presented in Table 2, and show that the probability of having 4 or 5 visits with gain or loss by chance alone is far less than observed, where 18 subjects were observed to have gain (7 expected by chance) and 16 were observed to have loss (<1 expected by chance). Furthermore, 16 and 17 subjects were observed to have gain or loss in 3 out of 5 yearly visits, yet only 14 and 8 were expected to have this result, respectively. We observed many fewer people than would be expected with 1 or 2 visits showing gain or loss, and more than expected with no gain or loss. This would indicate stability in MD scores among most patients.

We next investigated the relationship between IOP and VF gain or loss. A descriptive look at the average percent of visits in the first 5 years in which subjects showed a substantial gain (or loss) in MD from baseline, stratified by their maximum IOP during that time, showed an ordinal trend (Table 3). Specifically, subjects with the lowest maximum IOPs (< 13mmHg) had on average 18.7% of visits showing a substantial gain in MD compared to 13.4%, 10.6%, and 8.0% in subjects with maximum IOPs of 14–17mmHg, 18–21mmHg, and > 22mmHg. When looking at the percent of visits showing substantial loss, those subjects with lower maximum IOP had a smaller percentage of visits showing loss in MD than those with higher maximum IOP.

Results of a logistic regression that evaluated the association of baseline factors with VF improvement are shown in Table 4. The factors evaluated were the same as those previously tested for association with VF loss, as reported in Musch et al.<sup>7</sup> Main effects without interactions were found for gender (females were more likely to show VF improvement, OR=1.73; 95% CI, 1.17, 2.56) and an indicator variable for a visit conducted 1 year prior to cataract extraction (VF improvement was much less likely at this time, OR=0.11; 95% CI, 0.02, 0.62). A significant interaction between baseline MD and treatment is shown in Figure 2A, wherein participants treated with surgery who presented at baseline with more substantial VF loss were more likely to show VF improvement than those with comparable VF loss treated medically. An interaction of baseline vascular or cardiac disease (other than hypertension) with time (Figure 2B) indicated that participants with these conditions were increasingly less likely over time to show VF improvement than those who did not have these conditions.

To evaluate the effect of various IOP control measures on VF improvement, and deal with the substantial dependencies that exist among these various measures, we added each IOP control measure to the final multivariable logistic model independently and evaluated whether it contributed significantly to the model's fit. Table 4 provides the odds ratios and 95% confidence intervals for the IOP control measures, and Figure 3 displays a forest plot of the findings. Two continuous measures of IOP control during treatment were predictive of VF improvement: lower mean IOP (OR, 1.30; 95% CI, 1.10, 1.54) and lower minimum IOP (OR, 1.21; 95% CI, 1.04, 1.41), with both OR values relating to a 3.5 mmHg (one SD) decrease in these measures. For example, a person with a mean IOP decrease from baseline of 3.5 mmHg would have on average a 30% higher odds of VF improvement (OR=1.30). As the percent of IOP measurements <16, 18, or 20 mmHg increased (using follow-up data from 2 to 9 years after treatment initiation), the odds of substantial VF improvement increased (ORs=1.17, 1.20, and 1.30, respectively, for a 20% increase). The effects of IOP control measures were consistent in a model that included adjustment for center effects (results not shown).

Results from the Cox regression of time to sustained VF improvement were consistent with the logistic regression model results (Table 4) with respect to sex and cataract effects and most importantly, the IOP effects. All IOP effects were in the direction of better IOP control being associated with greater risk of VF improvement. Only the IOP effect for “All IOP <18 mmHg” was substantially different from the logistic model, with the Cox regression result showing a stronger relationship. In contrast to the logistic regression model, no interactions

were significant in the Cox model. This may be due to the reduced sample size over follow-up subsequent to improvement events, whereas in the repeated measures logistic model all follow-up data are used.

Finally, we present information on two CIGTS participants whose VFs demonstrated improvement. These subjects were selected from 12 patients who had VF improvement (at least a 3 dB gain from baseline) at every visit through 24 months (and were not missing any of those visits). Prior to treatment, these newly-diagnosed study participants' MD values were  $-7.79$  and  $-13.85$  dB, and they each improved by at least 3 dB after treatment. The first subject improved from a MD of  $-7.79$  dB to MDs measuring between  $-1.92$  and  $-3.54$  dB within the first 2 years post-surgery and showed continued improvement at 5 years with a MD of  $-5.05$  dB. The second subject entered the study with a MD of  $-13.85$  dB and improved to MDs between  $-10.28$  and  $-6.51$  dB over the first 2 years after surgery, and  $-4.27$  dB at 5 years. Of interest is the fact that both subjects had substantial IOP reduction after trabeculectomy that was maintained over 5 years of follow-up – one subject's study eye IOP decreased from a baseline of 44 mmHg to post-surgery IOP values consistently in a range from 6 to 11.5 mmHg, and the second subject's study eye IOP decreased from a baseline of 22 mmHg to post-surgery IOP values consistently in a range from 5 to 8.5 mmHg. A complete evaluation of all VFs for the temporal association between VF change and IOP level is beyond the scope of this paper, and will be the subject of further evaluation.

## Discussion

A long-held attitude among many ophthalmologists is that VF loss is due to disease progression and thereby is real, but VF gain cannot be real. We have demonstrated that the proportion with gain and loss is similar through five years after treatment initiation. We believe that approximately half of the observed improvement and loss reflects real VF change rather than variability of measurement. Our argument rests on consistency of findings over time and the strong association between improvement or loss and IOP control.

Several previous studies showed a quite substantial (2.8 dB) learning effect between first and second VF tests, but no significant change for subsequent VF tests.<sup>11,12</sup> All CIGTS subjects were required to have had a prior VF test. We reported an approximate 0.3 dB MD improvement in VF test results in CIGTS patients between the first and second baseline VF tests, with a diminution of this effect for tests conducted over the next 20 days.<sup>2</sup> Even the more prolonged increase in mean sensitivity for standard automated perimetry reported by Gardiner et al.<sup>13</sup> had a magnitude of 0.5 dB over the first year and then no significant change until after year 5 in subjects with suspected or early glaucoma. It would be difficult to attribute the lengthy pattern of substantial VF improvement we observed solely to a learning effect, as that would be expected to diminish over time with VF testing ongoing at six month intervals. The CIGTS protocol's requirements for prior VF experience also mitigate a learning effect.

Proper interpretation of the results from VF testing requires careful assessment of the role of noise. VF testing is subjective; results are influenced by patient characteristics such as attentiveness, fatigue, and anxiety. Variation will result from differences between how tests

are conducted by different technicians, or by the same technician at different times. Furthermore, as with any measurement, some level of inaccuracy is inevitable. We estimate that approximately 4% of gain or loss from baseline can be explained by variability in measurement. We instituted preventive measures in the CIGTS protocol to reduce this potential, including thorough training and certification of all test administrators. Even in the presence of variation, in most cases such noise should tend to decrease the ability to detect change over time, rather than produce a pattern of comparable increasing loss and improvement.

While the magnitude and pattern of VF improvement over time seem to point to this improvement representing real change, we looked at both descriptive information and the results of our statistical model to determine whether the observed improvement was real or spurious. The descriptive data (Table 3) reveal trends in both substantial VF loss and improvement by the extent to which IOP was controlled during treatment. Both trends are consistent with the concept that better IOP control is beneficial. The percentage of participants that had occurrences of VF improvement increased with better control of IOP, whereas substantial VF loss percentages increased with lesser control of IOP during treatment. If one assumes that VF results are random, there would be an equal amount of substantial VF loss and substantial VF improvement, as chance alone would dictate their frequency. This was not observed in our descriptive findings, as seen in the differences between percentages when these two outcomes are stratified by IOP control level.

The model-based findings of associations with indicators of better IOP control are particularly indicative of the observed VF improvement representing a real effect. VF improvement was more likely when the mean IOP over time was lower, when the minimum IOP was lower, and when three different measures of how frequently IOP values were less than a set value (16, 18, & 20 mmHg) increased. These IOP control measures remained significantly associated with VF improvement upon adjustment for baseline factors that were predictive of VF improvement. The interaction between baseline MD and treatment is also intriguing, as it demonstrates a predictive association of surgery with VF gain in those with more substantial VF loss at baseline. We know from our prior work that surgery reduced IOP more effectively than medicine and this IOP effect persisted over time, and that surgery in participants with more VF loss at baseline was associated with less VF loss than medical treatment.<sup>8,9</sup> These measures of treatment effect are all consistent with the VF improvement we observe being real. Our finding of female gender being associated with an increased likelihood of VF improvement, while deserving of further evaluation, lacks an evident rationale. Given this lack and the fact that we evaluated many covariates, a Type I error could underlie such a finding.

There is some evidence from past reports that substantiate VF improvement as a real possibility. Spaeth<sup>9</sup> analyzed information from 77 patients who underwent trabeculectomy and 195 patients receiving argon laser trabeculoplasty. Evaluating before and after (3 months and one year) measures of their VFs with Octopus perimetry, he reported a highly significant association of IOP reduction with VF improvement. In a subsequent report, Katz et al.<sup>14</sup> had three glaucoma specialists who were masked to the clinical course evaluate optic disc and VF findings before and after treatment. Their assessment yielded 31% (20/63) of



the VFs classified as improved, and this improvement was significantly associated with the extent of IOP reduction. Wilson et al.<sup>15</sup> reported that VF improvement occurred in 96% of 24 eyes that had primary open angle glaucoma and were treated by trabeculectomy. Finally, Parrish et al.<sup>16</sup> conducted a masked evaluation of change in optic disc cupping from baseline to five years in 348 study eyes of CIGTS participants. He reported a significant association ( $p < 0.001$ ) between reversal of optic disc cupping and lower post-trabeculectomy IOP. Although the group showing glaucomatous disc change ( $n=22$ ) had significantly more worsening of MD from baseline relative to both the no change group ( $n=298$ ) and the disc improvement group ( $n=22$ ), there was no significant difference in MD change between the latter two groups. We note that although MD provides a global assessment of VF change, it does not capture region-specific change and may mask changes in specific regions of the VF.

While regression to the mean (RTM) could be considered a possible cause of the observed improvement in VF, there are several reasons to believe that RTM is not a substantial contributor to these findings. First, we reduced measurement error by requiring multiple baseline measures of MD. The baseline value was calculated from the average of two, or median of three VFs, helping to reduce the effects of unreliable and unusually large or small measures. As unusually large differences in the MD values from three baseline VF tests are likely to be due to an anomalously low MD value (e.g., due to loss of attention or fixation), which would bias our results to show more improvement, we conducted a sensitivity analysis of our multivariable model in which those with more than a 3 dB difference in their baseline MD values were excluded ( $n=66$ ). The results were less significant due to lower power, but the significance of most effects was maintained, and the direction of all findings were consistent with the model reported in Table 3 (see supplementary Table 1; supplemental material at [AJO.com](http://AJO.com)). Second, we did not base enrollment on a level of MD indicative of substantial loss. Our eligibility criteria did not include a MD requirement. In fact, we enrolled some subjects (10.4%) whose MD was in a normal range (  $-1$ dB). Third, we did not conduct analyses based on any subset of subjects that was restricted to a worse baseline MD. Such subset analyses would be susceptible to RTM.<sup>17</sup> Fourth, a plot of MD change at 5 years by baseline MD did not show a pattern of improvement more distinct in those with worse baseline MD, indicative of RTM (see supplementary Figure 2; supplemental material at [AJO.com](http://AJO.com)). Lastly, we observed meaningful associations between VF improvement and measures of IOP control that would not be expected from random error associated with RTM.

In summary, we found a comparable percentage of CIGTS participants showed substantial (  $3$  dB) VF gains and losses over a five year period after treatment initiation. The potential that learning effects, variability in VF measurement, and/or chance could explain away VF gains was critically evaluated. Associations of VF gain with measures of sound IOP control during treatment, such as lower mean IOP and lower minimum IOP, as well as percentages of IOP measures lower than 16 mmHg, are consistent with VF gains being a real effect.

A recent report of work by Caprioli<sup>18</sup> notes that robust lowering of IOP not only slows the fast component of VF decay after surgery, but also results in a significant improvement of visual sensitivities at these VF locations. He characterizes this improvement as typical, not

uncommon, significant in amplitude, and it appears to be sustained. If damaged retinal ganglion cells whose function is being limited by high IOP or other glaucoma-induced effects can be revived by the dramatic lowering of IOP produced by treatment, which is especially a result of surgery, as first proposed almost 30 years ago,<sup>8</sup> one would expect in some cases to detect VF improvement in effectively treated subjects. We believe that such an effect is demonstrated in the VF data collected from CIGTS participants.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

### Funding/Support

This study was supported by a research grant from the NIH/NEI (#R21 EY020912), and by a departmental grant from Research to Prevent Blindness (RPB), New York, NY. Dr. Musch is a recipient of RPB's Lew R. Wasserman Merit Award.

### Financial Disclosures

DCM's institution received financial support for his participation in: (a) the CIGTS (NIH grant EY09148) (b) this work (NIH grant EY020912), (c) a planning grant for a multicenter trial of statin treatment for glaucoma (NIH grant EY022399), and (d) from Pfizer to study the cost of visual field progression. He is a data and safety monitoring board member for Ivantis and a consultant to Glaukos and InnFocus. He is a co-submitter of a US patent application (now pending) of an algorithm that allows optimizing the frequency of follow-up in patients diagnosed with glaucoma.

PFP received travel support and honoraria for participation as a member of the DSMC for the CIGTS, he is a board member for Aeon Astron, a consultant to Aurolab & Aravind Eye Hospitals, AqueSys, and InnFocus, and has received payment for lectures from AqueSys and InnFocus.

GS received grant money to his institution from Merck to study quality of life, received payment for lectures from Merck and Allergan, received royalties from Elsevier for surgical text, and received payment for development of educational presentations from Merck.

BWG, LMN, and PRL's institution received financial support for their participation in this work (NEI grant EY020912), and BWG and PRL's institution received financial support for their participation in the CIGTS (NIH grant EY09148).

### Other Acknowledgments

The authors would like to thank Kyoko Ishida, MD, PhD, of Gifu University, whose previous work with Dr. Paul Palmberg on visual field improvement and its association with IOP control contributed to the initiative of this study.

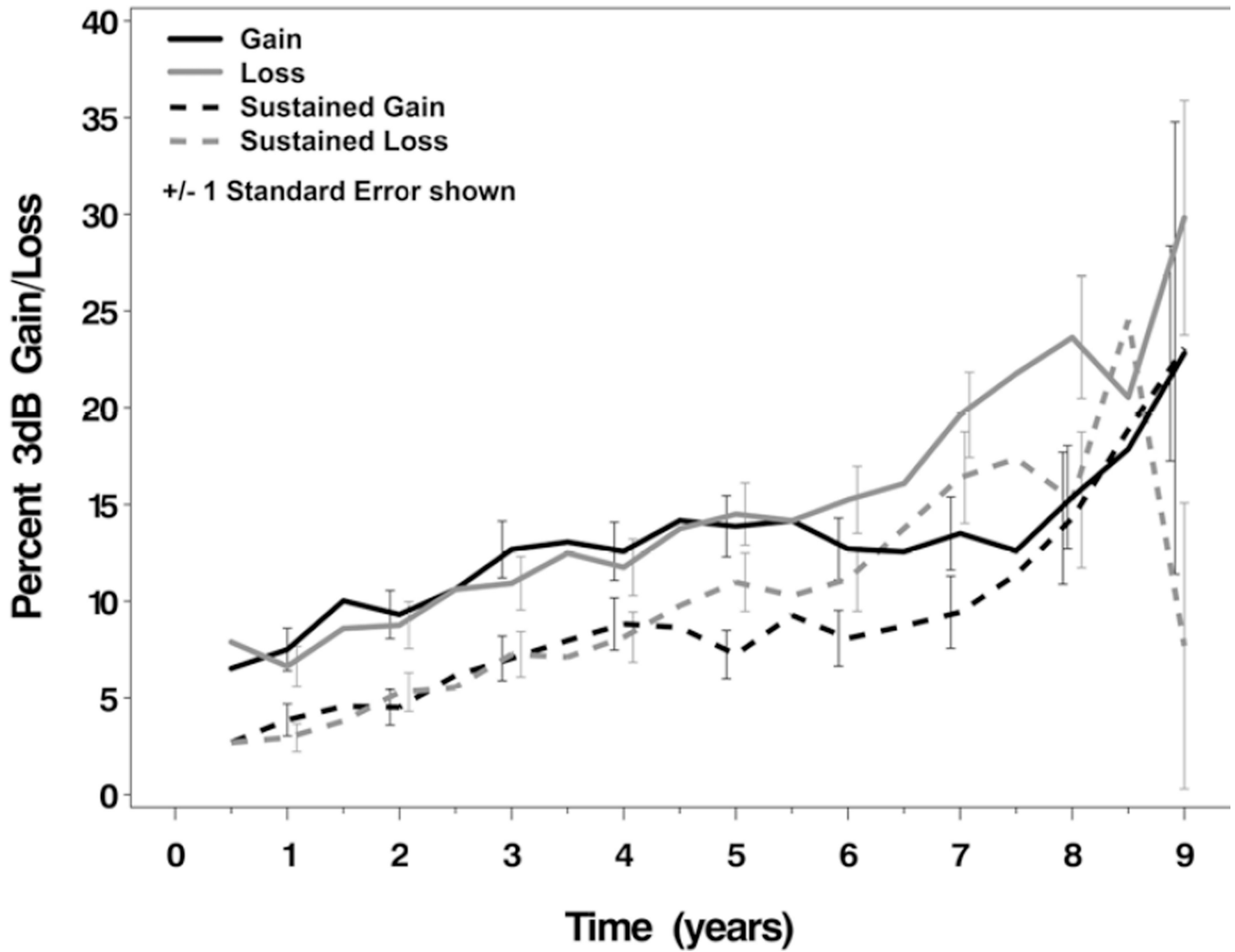
This paper represents a secondary data analysis from a randomized clinical trial (CIGTS) of initial treatment of glaucoma with medications versus trabeculectomy. The trial started in 1993 and was registered at [www.clinicaltrials.gov](http://www.clinicaltrials.gov) on September 23, 1999, registration number NCT00000149

A list of the CIGTS investigators, clinical and resource centers, and the data and safety monitoring committee can be found in Appendix 1 of Musch DC et al., *Ophthalmology* 2009; 116(2):200–207.

## References

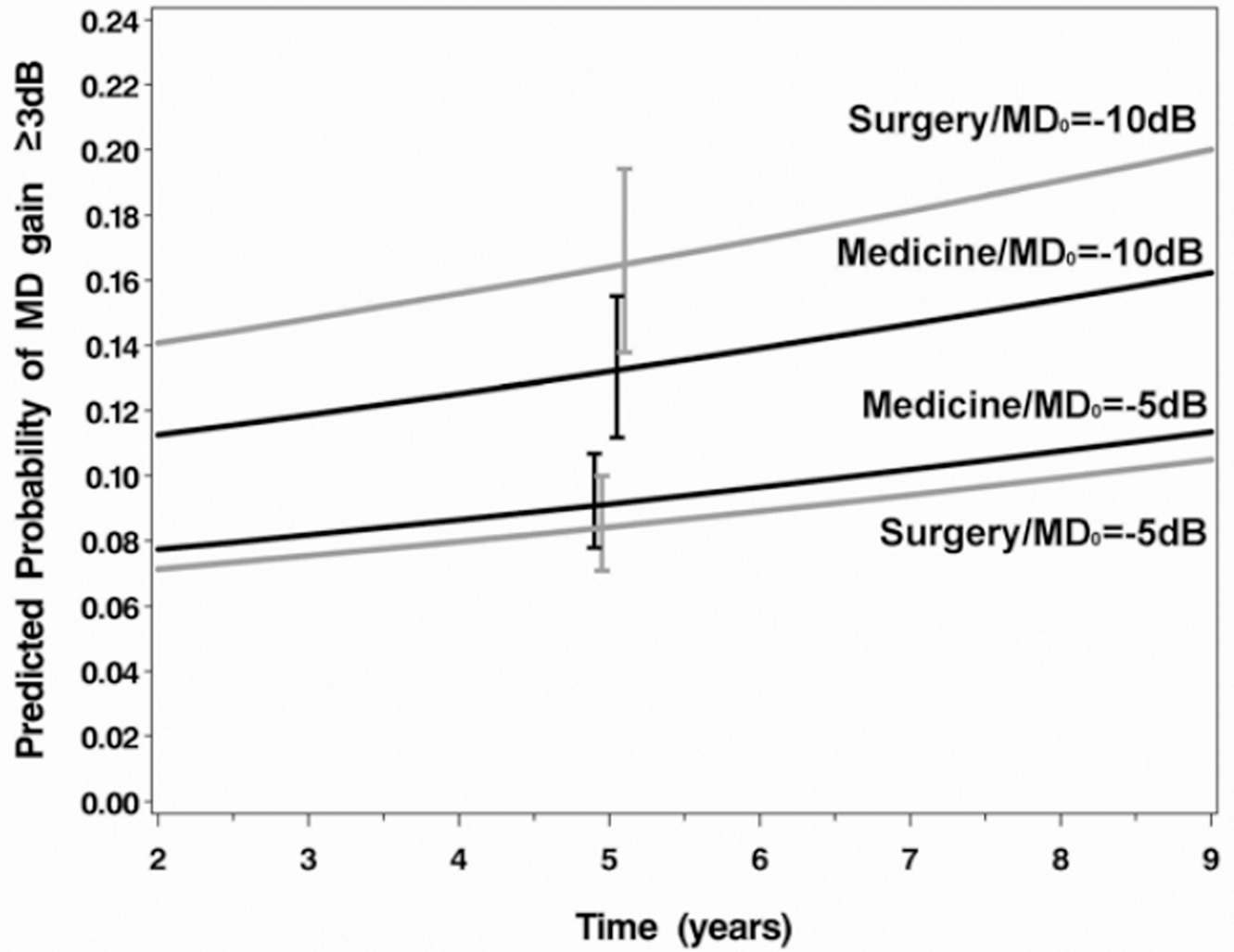
1. Spry PGD, Johnson CA, McKendrick AM, Turpin A. Measurement error of visual field tests in glaucoma. *Br J Ophthalmol.* 2003; 87(1):107–112. [PubMed: 12488273]
2. Gillespie BW, Musch DC, Guire KE, et al. The Collaborative Initial Glaucoma Treatment Study: baseline visual field and test-retest variability. *Invest Ophthalmol Vis Sci.* 2003; 44(6):2613–2620. [PubMed: 12766064]

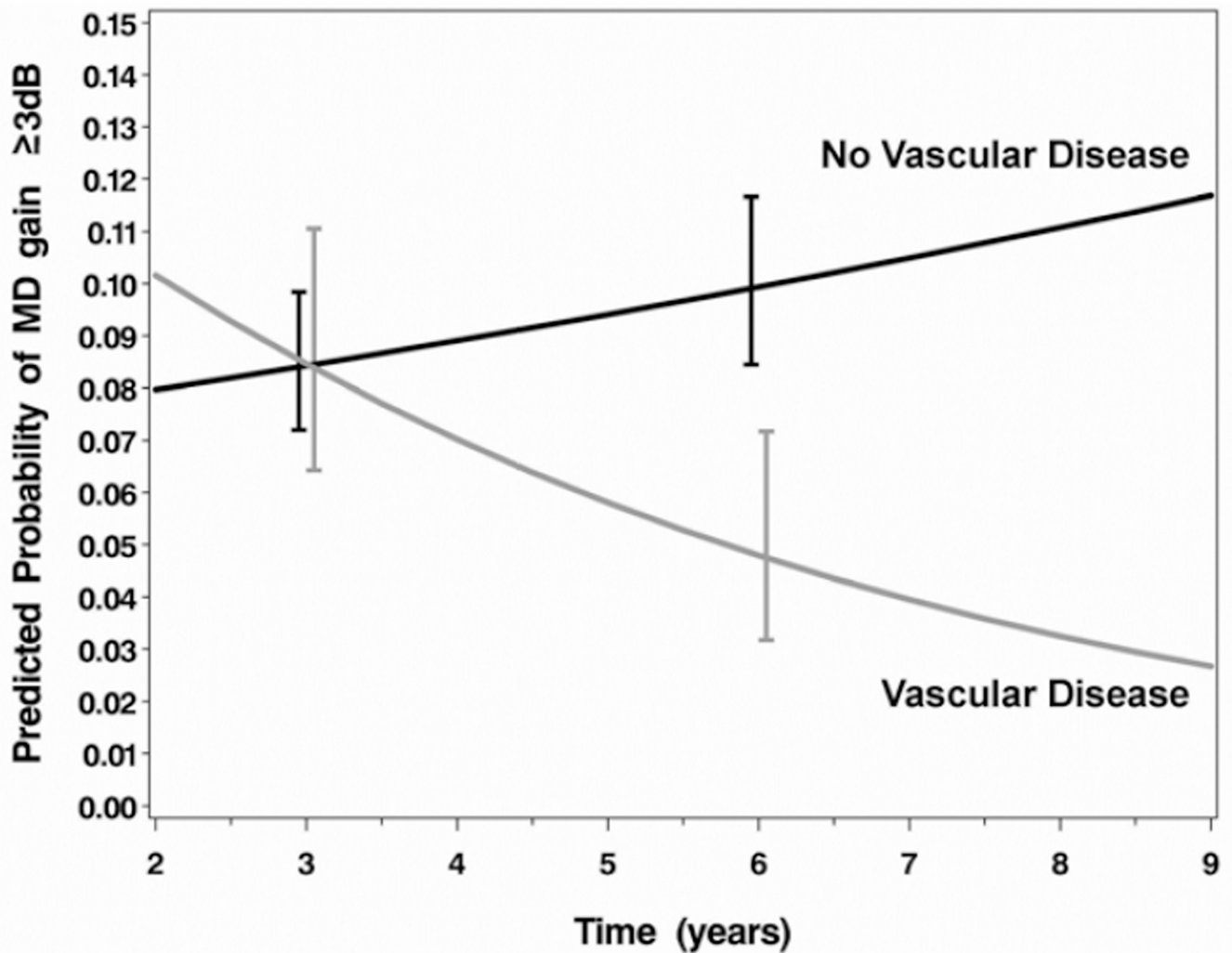
3. Johnson CA. Standardizing the measurement of visual fields for clinical research. Guidelines from the Eye Care Technology Forum. *Ophthalmology*. 1996; 103(1):186–189. [PubMed: 8628553]
4. The AGIS Investigators. The Advanced Glaucoma Intervention Study (AGIS): 1. Study design and methods and baseline characteristics of study patients. *Control Clin Trials*. 1994; 15(4):299–325. [PubMed: 7956270]
5. Musch DC, Lichter PR, Guire KE, Standardi CL. CIGTS Study Group. The Collaborative Initial Glaucoma Treatment Study. Study design, methods, and baseline characteristics of enrolled patients. *Ophthalmology*. 1999; 106(4):653–662. [PubMed: 10201583]
6. Gordon MO, Kass MA. Ocular Hypertension Treatment Study Group. The Ocular Hypertension Treatment Study. Design and baseline description of the participants. *Arch Ophthalmol*. 1999; 117(5):573–583. [PubMed: 10326953]
7. Musch DC, Gillespie BW, Lichter PR, Niziol LM, Janz NK. CIGTS Study Investigators. Visual field progression in the Collaborative Initial Glaucoma Treatment Study. The impact of treatment and other baseline factors. *Ophthalmology*. 2009; 116(2):200–207. [PubMed: 19019444]
8. Musch DC, Gillespie BW, Niziol LM, Lichter PR, Varma R. CIGTS Study Group. Intraocular pressure control and long-term visual field loss in the Collaborative Initial Glaucoma Treatment Study. *Ophthalmology*. 2011; 118(9):1766–1773. [PubMed: 21600658]
9. Spaeth GL. The effect of change in intraocular pressure on the natural history of glaucoma: lowering intraocular pressure in glaucoma can result in improvement of visual fields. *Trans Ophthalmol Soc UK*. 1985; 104(3):256–264. [PubMed: 3861005]
10. Zeger SL, Liang KY, Albert PS. Models for longitudinal data: a generalized estimating equation approach. *Biometrics*. 1988; 44(4):1049–1060. [PubMed: 3233245]
11. Heijl A, Bengtsson B. The effect of perimetric experience in patients with glaucoma. *Arch Ophthalmol*. 1996; 114(1):19–22. [PubMed: 8540846]
12. Wild JM, Haarle AET, Dengler-Harles M, O’Neill EC. Long-term follow-up of baseline learning and fatigue effects in the automated perimetry of glaucoma and ocular hypertension. *Acta Ophthalmol (Copenh)*. 1991; 69(2):210–216. [PubMed: 1872140]
13. Gardiner SK, Demirel S, Johnson CA. Is there evidence for continued learning over multiple years in perimetry? *Optom Vis Sci*. 2008; 85(11):1043–1048. [PubMed: 18981919]
14. Katz LJ, Spaeth GL, Cantor LB, Poryzees EM, Steinmann WC. Reversible optic disk cupping and visual field improvement in adults with glaucoma. *Am J Ophthalmol*. 1989; 107(5):485–492. [PubMed: 2712131]
15. Wilson MR, Kosoko O, Cowan CL Jr, et al. Progression of visual field loss in untreated glaucoma patients and glaucoma suspects in St. Lucia, West Indies. *Am J Ophthalmol*. 2002; 134(3):399–405. [PubMed: 12208252]
16. Parrish RK II, Feuer WJ, Schiffman JC, et al. Five-year follow-up optic disc findings of the Collaborative Initial Glaucoma Treatment Study. *Am J Ophthalmol*. 2009; 147(4):717–724. [PubMed: 19152871]
17. Barnett AG, van der Pols JC, Dobson AJ. Regression to the mean: what it is and how to deal with it. *Int J Epidemiol*. 2005; 34(1):215–220. [PubMed: 15333621]
18. Caprioli J. Glaucoma: a disease of early cellular senescence. *Invest Ophthalmol Vis Sci*. 2013; 54:ORSF60–ORSF66. [PubMed: 24335071]



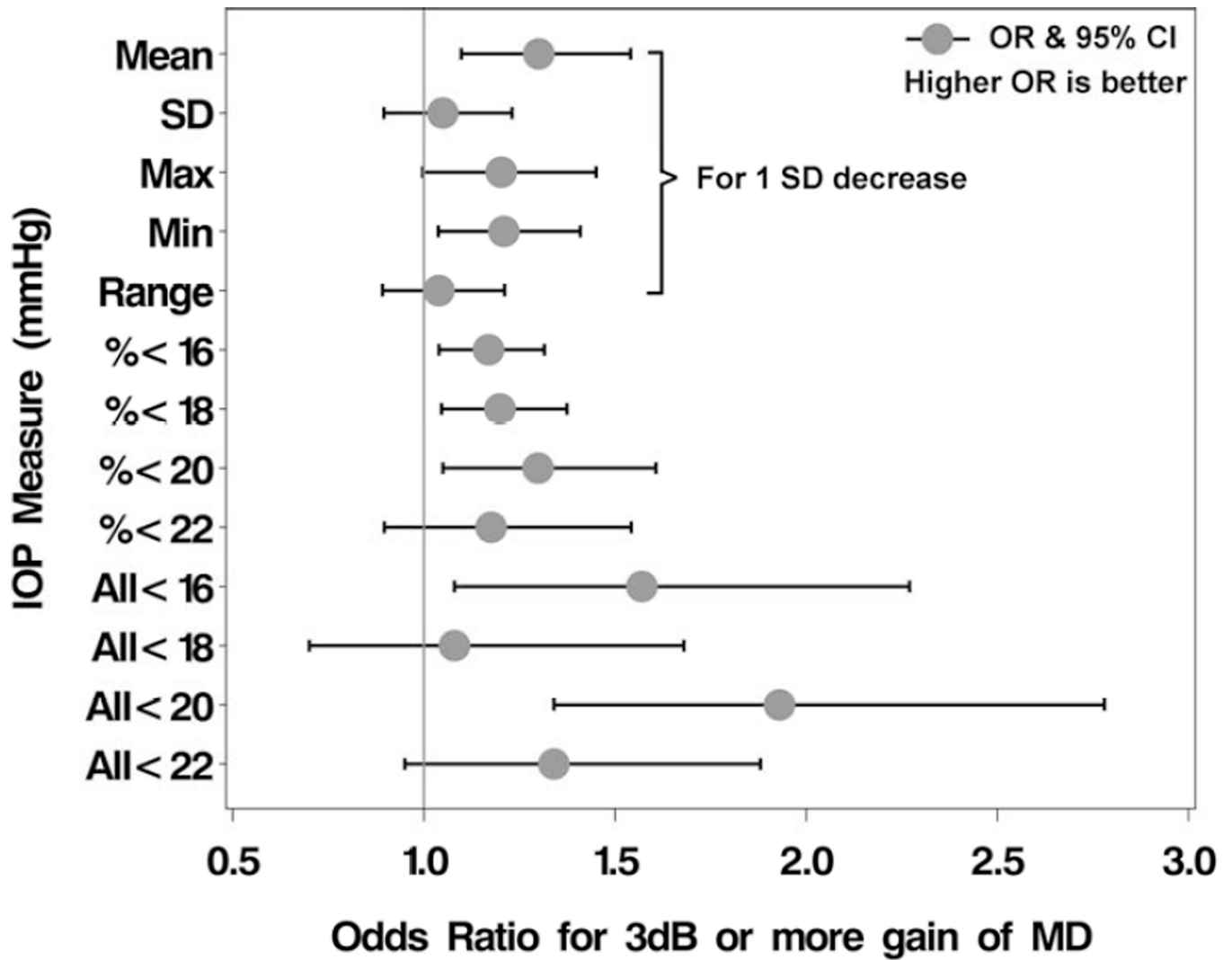
**Figure 1.**

Percent of subjects showing a substantial gain/loss ( $\geq 3$  dB) in mean deviation from baseline (evaluated at each visit) and percent of subjects showing sustained gain/loss (validated by the next consecutive visit) over 9 years of follow-up in the Collaborative Initial Glaucoma Treatment Study. Standard error bars were calculated assuming estimates from a binomial distribution. Percent gain was significantly different from percent loss at years 7 and 8 (p-values=0.0053 and 0.0088, respectively; test of equality of binomial proportions). dB=decibel.





**Figure 2.** Interaction plots for the effects of (A) baseline mean deviation and treatment on substantial visual field improvement (increase in mean deviation (MD) of  $\geq 3$ dB from baseline) and (B) vascular/cardiac disease over time on substantial visual field improvement in the Collaborative Initial Glaucoma Treatment Study. Error bars are  $\pm 1$  standard error. dB=decibel.



**Figure 3.** Association of intraocular pressure control measures with substantial visual field improvement (increase in mean deviation (MD) of 3dB from baseline) in the Collaborative Initial Glaucoma Treatment Study. OR=odds ratio; CI=confidence interval; SD=standard deviation; IOP=intraocular pressure; mmHg=millimeters of mercury; dB=decibel.

**Table 1**

Descriptive statistics of the Collaborative Initial Glaucoma Treatment Study sample at baseline, including patient demographics and clinical characteristics.

Continuous Variables	N	Mean(SD)	Min, Max	Median
Follow-up (years)	607	7.2 (2.3)	0.0, 14.5	7.7
Age (years)	607	58.0 (10.9)	28.8, 75.8	59.2
MD (decibels)	607	-5.4 (4.3)	-23.5, 3.4	-4.4
IOP (mmHg)	607	27.5 (5.6)	19.0, 50.0	27.0
VA (letters)	607	85.7 (5.7)	70.0, 99.0	86.0
Categorical Variables	Frequency (Percent)			
Sex				
Female	273 (45.0)			
Male	334 (55.0)			
Race				
White	337 (55.5)			
Black	231 (38.1)			
Asian	10 (1.7)			
Other	29 (4.8)			
Education				
<High School	128 (21.1)			
High School	167 (27.5)			
>High School	312 (51.4)			
Diagnosis				
POAG	550 (90.6)			
Pseudoexfoliation	29 (4.8)			
Pigmentary	28 (4.6)			
Diabetes	102 (16.8)			
Hypertension	225 (37.1)			
Other Vascular/Cardiac Disease	91 (15.0)			
Smoking Status				
Never smoker	234 (38.6)			
Ex-smoker	246 (40.5)			
Current smoker	127 (20.9)			
Immediate Family Hx (n=545)	201 (36.9)			
Distant Family Hx (n=461)	112 (24.3)			

MD=mean deviation; IOP=intraocular pressure; VA=visual acuity; SD=standard deviation; Hx=history.



**Table 2**

Observed and expected occurrences of  $\geq 3$  dB visual field gain or loss from the baseline mean deviation over fiveannual follow-up visits in the Collaborative Initial Glaucoma Treatment Study (CIGTS). Expected occurrences were simulated assuming no visual field change over time, i.e., observed gains or losses were due to chance alone.\*

# visits with $\geq 3$ dB gain or loss	Gain		Loss	
	Observed	Simulated Expected	Observed	Simulated Expected
0	306	282.7	312	244.8
1	45	76.1	46	123.5
2	28	32.5	22	35.3
3	16	14.4	17	8.1
4	12	6.2	12	1.2
5	6	1.20	4	0.04

dB=decibel; MD=mean deviation

\* Only CIGTS cases with complete data for annual visits 1–5 were included; the intervening 6 month visits were excluded to increase the sample size with complete data (n=413). We calculated the expected number of yearly visits with gain (0,1, 2, 3, 4, or 5), assuming baseline and follow-up MDs had the same marginal mean ( $-5.4$  dB) and standard deviation (4.3 dB). Data were simulated assuming multivariate normality on a transformed scale ( $\ln(-MD+5)$ ) to reduce skewness, with correlation structure observed in the CIGTS data:  $r=0.90$  between the two baseline measures;  $r=0.85$  between the mean baseline measure and any of the follow-up measures. Data were back-transformed to the raw scale to assess gain and loss. The asymmetrical transformation resulted in minor differences between the expected results for gain and loss. We calculated the number of subjects expected to show gain by chance as the proportion of 10,000 simulated patients in which the VF gain, compared to the average of the two baseline measures, was at least 3 dB. These proportions were multiplied by the sample size (413) to yield the expected number of CIGTS subjects with gain by chance alone. Calculations for loss were similar.

**Table 3**

Descriptive statistics on the percent of visits during the first 5 years showing a 3 decibel improvement or loss of mean deviation from baseline, stratified by maximum intraocular pressure in the first 5 years, in the Collaborative Initial Glaucoma Treatment Study. Statistics were calculated on the subset of subjects who had at least 5 years of follow-up and did not miss their 5 year visit (n=476).

	% visits 3dB gain	% visits 3dB loss	% Difference	MD change @ 5 years
<b>Maximum IOP in first 5 years</b>	<b>Mean (SD)</b>			
13 mmHg (n=24)	18.7 (30.0)	6.3 (14.0)	12.5 (36.2)	0.39 (5.38)
14–17 mmHg (n=80)	13.4 (24.4)	6.3 (16.1)	7.1 (31.6)	0.65 (2.79)
18–21 mmHg (n=151)	10.6 (19.0)	11.7 (20.6)	–1.2 (31.2)	0.01 (3.60)
22 mmHg (n=221)	8.0 (17.1)	10.6 (19.1)	–2.5 (27.4)	–0.46 (3.67)

Note: difference was calculated as % gain - %loss; dB=decibel; MD=mean deviation; IOP=intraocular pressure; SD=standard deviation; mmHg=millimeters of mercury; MD change was calculated as 5 year measure – baseline measure; the maximum number of visits for an individual over 5 years follow-up is 11 visits; 33.8% of study participants had at least 1 visit with gain and 34.5% had at least 1 visit with loss in the first 5 years.

**Table 4**

Model results for associations with substantial visual field improvement (  $\geq 3$  decibels improvement in mean deviation from baseline measurement) and sustained visual field improvement (  $\geq 3$  decibels improvement from baseline validated by at least 2 consecutive visits showing improvement) in the Collaborative Initial Glaucoma Treatment Study.

Main Effects	Repeated Measures Logistic Regression of Visit-specific Visual Field Improvement			Cox Regression of Time to Sustained (for 2 consecutive visits) Visual Field Improvement		
	OR	95% CI	P-value <sup>b</sup>	HR	95% CI	P-value <sup>b</sup>
Sex (Female vs. Male)	1.73	(1.17, 2.56)	0.0061	1.77	(1.20, 2.61)	0.0039
Cataract <sup>d</sup>	0.11	(0.02, 0.62)	0.0127	0.45	(0.06, 3.25)	0.4304
Baseline MD (per dB)	See interaction with treatment					
<b>Interactions</b>						
Treatment*Baseline MD			0.0403			
Surgery vs. Medicine @ MD <sub>0</sub> =-5dB	0.92	(0.62, 1.34)				
Surgery vs. Medicine @ MD <sub>0</sub> =-10dB	1.29	(0.81, 2.05)				
Surgery vs. Medicine @ MD <sub>0</sub> =-15dB	1.82	(0.90, 3.89)				
Other Vascular Disease*Time			0.0136			
Vascular Disease vs. None @ 3 years	1.00	(0.57, 1.77)				
Vascular Disease vs. None @ 5 years	0.59	(0.30, 1.19)				
Vascular Disease vs. None @ 7 years	0.35	(0.13, 0.96)				
<b>IOP Variables<sup>e</sup></b>						
Mean IOP <sup>d</sup>	1.30	(1.10, 1.54)	0.0024	1.23	(1.05, 1.43)	0.0086
Standard Deviation IOP <sup>d</sup>	1.05	(0.90, 1.23)	0.5520	0.99	(0.83, 1.18)	0.9048
Maximum IOP <sup>d</sup>	1.20	(1.00, 1.45)	0.0540	1.22	(1.00, 1.48)	0.0467
Minimum IOP <sup>d</sup>	1.21	(1.04, 1.41)	0.0149	1.21	(1.04, 1.42)	0.0148
Range IOP <sup>d</sup>	1.04	(0.89, 1.21)	0.6225	1.02	(0.82, 1.28)	0.8286
% IOP < 16 mmHg <sup>e</sup>	1.17	(1.04, 1.31)	0.0092	1.13	(1.02, 1.24)	0.0205
% IOP < 18 mmHg <sup>e</sup>	1.20	(1.05, 1.37)	0.0091	1.18	(1.04, 1.33)	0.0091
% IOP < 20 mmHg <sup>e</sup>	1.30	(1.05, 1.61)	0.0162	1.18	(1.00, 1.39)	0.0501

No significant interactions

Main Effects	Repeated Measures Logistic Regression of Visit-specific Visual Field Improvement		Cox Regression of Time to Sustained (for 2 consecutive visits) Visual Field Improvement			
	OR	95% CI	P-value <sup>b</sup>	HR	95% CI	P-value <sup>b</sup>
% IOP < 22 mmHg <sup>e</sup>	1.18	(0.90, 1.54)	0.2421	1.15	(0.93, 1.41)	0.2055
All IOP < 16 mmHg	1.57	(1.08, 2.27)	0.0172	1.41	(0.93, 2.13)	0.1066
All IOP < 18 mmHg	1.08	(0.70, 1.68)	0.7164	1.53	(1.03, 2.26)	0.0345
All IOP < 20 mmHg	1.93	(1.34, 2.78)	0.0004	1.65	(1.09, 2.51)	0.0188
All IOP < 22 mmHg	1.34	(0.95, 1.88)	0.0922	1.36	(0.86, 2.16)	0.1890

OR=Odds ratio, CI=Confidence interval, HR=Hazard ratio, MD=Mean deviation, IOP=Intraocular pressure, mmHg=millimeters of mercury;

<sup>a</sup> Indicator variable for a visit 1-year prior to cataract extraction, a non-significant cataract effect was retained in the Cox model due to its strong hazard ratio and because it has been a standard adjustment in most prior CIGTS analyses;

<sup>b</sup> P-values are based on 1 degree of freedom Wald tests from the logistic regression or Cox regression;

<sup>c</sup> IOP variables were added individually to the model with main effects and interaction;

<sup>d</sup> ORs/HRs given for a 1 SD decrease (3.5 mmHg for mean, 1.5 mmHg for SD, 5.5 mmHg for Max, 3.5 mmHg for Min, 4.5 mmHg for Range);

<sup>e</sup> ORs/HRs given for a 20% increase.