STATISTICAL ANALYSIS

Protocol #CVI08008 (MIST401) V3.0 NCT01729208

April 30, 2018

Sample Size Estimation

The protocol originally estimated a sample size to detect a refractive treatment effect of 0.75D over 3 years. The target effect for sample size calculation was conservatively stated as "0.25D per year (i.e. 0.75D over 3 years". Using 0.25D per year as a target for sample size, it was estimated that 87 subjects per group would be needed (two-sample t-test with equal variance, P=0.05, power=90%). The protocol anticipated an enrolment target of 150 eligible subjects per group to account for attrition (14%/year) over the 3-year period.

Due to a longer than expected recruitment period it became evident that the number of subjects enrolled would be smaller than this target. A review of the sample size revealed that the actual number of subjects enrolled would be fully sufficient to detect the difference of 0.75D over 3 years.

Statistical Methods

Analysis Data Set

Although the protocol stated that the per-protocol population would be used for the primary analysis, an intent-to-treat was later felt to be more appropriate.

All available data from all subjects were used in the statistical analyses.

Missing Data – Primary Outcomes

No adjustments were made for missing data in the main analyses. ███████████ ███ Only one exclusion from the effectiveness analysis was made at the 36-month visit for a MiSight subject that had begun growth hormone therapy six months earlier.

An additional sensitivity analysis was made to estimate the impact of the missing effectiveness data. The primary effectiveness outcomes were change from baseline in cycloplegic SERE and axial length measured at annual visits. For subjects that discontinued prior to the first annual visit, no information regarding their myopic progression rate could be made and no imputation was felt appropriate. For subjects with 12-month data only, the same progression rate was assumed in the analysis for the 2nd and 3rd years. For subjects with 12-and 24-month data only, the highest rate of these two years was assumed for the 3rd year.

Data Pooling

Data were pooled from multiple study sites for this analysis with the justification for pooling based on three factors:

- The study sites implemented one common protocol.
- The CRO has closely monitored study site protocol compliance.
- The study sites used common data collection procedures.

In addition, as specified in the protocol, an analysis of the interaction of lens type by site was performed and found not significant ($p > 0.10$) for the primary effectiveness variables.

<u>Analysis of Baseline Data</u>

During the enrolment phase of the study, the comparability between test and control groups was evaluated by the two-sample t-test (continuous data), Mann-Whitney U-test (categorical data) or Fisher's exact test (nominal data). Imbalances of potentially confounding variables that were identified between the two groups were to be addressed by including them as covariates in the final analysis.

All inference was carried out with the type I error rate controlled at 5%.

<u>Primary Safety Analysis</u>

Comparisons between the lenses were to be made using two-sided 95% confidence intervals of odds ratios (test over control). Non-inferiority would be concluded if the upper limit of the confidence limit is less than 1.33.

<u>Primary Effectiveness Analysis</u>

Unadjusted mean change in refractive error and axial length were checked for normality and compared between treatment groups using a Student t-test.

These primary outcome measures were also compared using linear mixed models. The linear-mixed model included treatment, visit, site, and all two- and three-way interactions as fixed effects; subject age group, sex, ethnicity, and baseline myopia (refractive error or axial length) as fixed covariates; and subject (nested in site) and eye as random effects.[35] This model differed slightly from that specified in the protocol where site was specified as a random factor.

Comparisons between the test and control lenses were carried out using a two-sided confidence interval constructed least-square mean differences at each follow-up visit. A statistical difference was concluded if the 95% confidence limit of the mean difference is greater than 0 (test minus control).

At the time of the protocol, we suggested that a clinically significant difference would be concluded if the lower 95% confidence limit was ³0.25D/year (0.75D total) after 3 years. Although this was not achieved, a significant clinical benefit has been demonstrated ▮▮▮▮▮▮▮

<u>Statistical Software</u>
The statistical analyses were undertaken using SAS version 9.4.