Clinical Investigation of the VR-800 Device:

Statistical Methods and Analytic Considerations

NCT Number: NCT04208750

Document Date: 03 FEB 2020

Statistical methods and analytic considerations

Sample size. An initial calculation was made for the continuous subjective rating data to determine the sample size required to detect an expected small effect size (0.3) using a paired samples-t-test targeting 80% power and level of significance of 5% (two-sided). This calculation indicated that a total sample size N=90 was required to detect an effect size of 0.3 in the paired mean difference between the Vision-R^{TM} 800 and the standard refractions. Because several of our subjective measures were based on a comparison of proportions in the full sample (N=90) we also estimated the sample size needed for testing a difference in proportions of Δp = 0.65-0.35 = 0.3 in a paired-sample design. We found that the required sample size of N=85 would achieve 80% power at the two-sided 5% level of significance.

To allow for an attrition rate of 25%, we increased the larger of the two sample size estimates (N=90) to obtain a post-screening study sample size of N=120.

Overview. All analyses were performed using R v3.5.3 (The R Foundation; www.r-project.org) and SAS v9.4 (SAS Institute, Cary, NC, USA). For each outcome measure, all available data were incorporated into the statistical analyses and modeling. Although not a primary comparison, habitual lens information is included in some instances to provide additional context for comparison of the Vision-R^{TM} 800 and standard refraction technologies data.

Refractive error. Clinically obtained measures of refractive error measures (sphere, cylinder and axis) were converted to Cartesian power vector coordinates (M, J0, and J45)(Thibos, Wheeler et al. 1997, Thibos and Horner 2001) and refractive error differences depicted graphically in three-dimensional plots, conditioned by both refraction technology and participant group. These data were summarized in terms of parameter estimates of location (i.e., mean and median) and dispersion (i.e., standard deviation, standard error of the mean, median, minimum, maximum, range and 95% confidence intervals based on quantiles of the Student's *t*-distribution). Multivariate tests were used to statistically evaluate hypotheses related to differences between the refraction technologies.

LogMAR visual acuity. Sample distributions of logMAR visual acuity for each refraction method conditioned by participant group and viewing distance were depicted graphically using box-and-whisker plots and analyzed using linear mixed-effects (LME) regression. Separate models were fit to the near and distance data by restricted maximum likelihood estimation. Each LME model included fixed-effects terms for refraction method, participant group, and the refraction x participant group interaction as well as participant-specific random intercepts.

Subjective preferences and rating-scale data. All categorical measures were summarized and depicted graphically in terms of percentages of participants endorsing each response category.in For descriptive purposes, we report both a point estimate for each response percentage along with a 95% confidence interval to quantify the precision of the estimate. Bayesian methods were used to obtain estimates of [1] the relative frequency of participants' preferences for, confidence in, and perceived benefits of the Vision-R^{TM} 800 refraction technology and [2] the Bayes factor that quantifies the relative probability of the observed data under two competing models (e.g., null and alternative hypotheses). Freidman rank sum tests with pairwise post-hoc tests were used to analyze symptom frequency ratings for habitual, Vision-R^{TM} 800, and standard glasses.

Missingness due to COVID-19. There were two main data-related issues in this study due to the COVID-19 global pandemic and subsequent lockdown that began while the study was underway. First, because the pandemic prevented many participants from attending their final in-office visit, final subjective questionnaire responses were collected for some participants using an electronic survey. Post-hoc analyses indicated no significant differences between measures collected in-office and those collected remotely. Thus, all in-office and remotely collected data were pooled for purposes of analysis. Second, we note that differences in the final sample size of the two study groups (N=49 versus N=39 for young adults and presbyopes, respectively) are related to recruitment process at the beginning of the study rather differential drop-out during the study. Third, statistical comparisons between response proportions for the refraction technologies in the full sample remains appropriately powered. However, for conditional analyses comparing refraction methods within the young adult and presbyope subgroups, some caution is advised as the implied power is slightly under the targeted level of 80% for intermediate effect sizes.