**Pilot study**

**SD-OCT Image Variability Considerations**

We conducted a pilot study of the SD-OCT methodology to measure RNFL thickness, optic nerve head (ONH) volume, total retinal thickness (TRT), and ONH elevation the IIHTT OCT substudy. We assessed the variability for each SD-OCT parameter and the two approaches to segmentation analysis of the structure changes resulting from swelling of the ONH and peripapillary retina. Four sites including (State University of New York at Stony Brook, the University of Iowa, the University of California, Davis, and Roosevelt Hospital) collected the optic disc 200 x 200 volume scan and HD 5-Line Raster scan (centered on the optic disc) data using the same model SD-OCT machine and software (Cirrus 4000, Carl Zeiss Meditec, Inc., Dublin, CA, 6.01 software version). Men and women with one or more swollen optic discs (defined as thicker than the 95th percentile of normal average RNFL; 110 μm for 32 year olds), within the age range: 18 – 80 years who were undergoing SD-OCT for his/her disease management were eligible for the pilot study. ONH swelling could be due to papilledema, anterior ischemic optic neuropathy (AION), or optic neuritis. Subjects with optic nerve drusen or pseudopapilledema were excluded.

Each subject had study images collected on only one visit. OCT data from 18 eyes with ONH swelling from 14 subjects (one eye from 10 subjects and both eyes from 4 subjects) were collected. Following pupillary dilation with tropicamide, the seated subject was required to look into the SD-OCT instrument. Each subject had repeat SD-OCT images collected three times after getting up and being reseated/repositioned at the SD-OCT machine in order to test repositioning effects on the measurement variability of image collection and analysis for each affected eye.
Data were analyzed by (1) using repeated measure models to account for correlation of observations from the same subject and differences across images, (2) transforming volume, total retinal thickness and ONH height rating to z-scores by subtracting mean across all images and subjects and dividing this difference between an individual eye’s value and the mean by the standard deviation of all of the eyes, (3) using repeated measures models to analyze the z-scored volumes, thicknesses and raster ONH maximum height value; the within-person variability was estimated and then compared, and (4) computing within-person coefficient of variation for ONH height, optic nerve head volume, RNFL and total retinal thickness. The z-scores were used to represent the data from the 18 eyes on the same scale. For two readers, reliability (test/retest) was assessed by calculating intraclass correlation correlations (ICC) for each of the 5 images collected across the optic disc from the 5-line raster method, separately using only one eye per subject. Using repeated measures models to account for correlation of observations from the same subject, differences between readers were assessed. The differences were calculated across the images when average ratings between the two raters were used.

Pilot study results.

Using one eye per person, there were no significant differences across repeatedly collected images for RNFL thickness (p > 0.40), optic nerve head volume (p>0.50), TRT (p > 0.90), or raster ONH height (p = 0.2). The results were similar if data from both eyes were used (accounting for the inter-eye correlation). For RNFL thickness, ONH volume, and TRT, it was not necessary to transform the values to standard deviations and z-scores as was needed for ONH height values, in order to compare reproducibility of the different parameters for each eye. When comparing reproducibility, the within-subject standard deviation for ONH volume was 0.12 and the within-subject standard deviation for average raster ONH height for two readers was 0.20.
The within-person coefficient of variation (standard deviation/mean) was 0.057 for RNFL, 0.01 for ONH volume, 0.01 for total retinal thickness, and 0.036 for ONH height. Values were similar when measured from different images for each parameter. There was no difference by reader for the ONH height measurements, when looking across all images and subjects, using one eye per person for the raster scans ($p=0.69$). If both eyes were utilized per subject (but still accounting for the inter-eye correlation), results were similar. Intraclass correlations (ICC) were all very high (0.99) across each of the images, thus readers were very close in their ratings. The analyses showed reproducibility of RNFL thickness, ONH volume, TRT, and manual raster ONH height measurements.