



Published in final edited form as:

Ophthalmology. 2012 December ; 119(12): 2549–2557. doi:10.1016/j.ophtha.2012.06.040.

Optical Coherence Tomography grading reproducibility during the Comparison of Age-Related Macular Degeneration Treatments Trials

Francis Char DeCroos, MD^{1,2}, Cynthia A. Toth, MD², Sandra S. Stinnett, DrPH², Cynthia S. Heydary, BA², Russell Burns, AAS², Glenn J. Jaffe, MD², and for the CATT Research Group

¹Wills Eye Institute (Mid Atlantic Retina), Philadelphia, PA 19107

²Duke University Eye Center, Durham, NC 27707

Abstract

OBJECTIVE—To report Reading Center reproducibility during grading of Optical Coherence Tomography (OCT, Carl Zeiss Meditec, Dublin, California) images obtained during the Comparison of Age-Related Macular Degeneration Treatments Trials (CATT).

DESIGN—Prospective clinical trial

PARTICIPANTS—Independent reading teams reevaluated 270 OCT scans randomly sampled from the first 2 years of CATT enrollment. To assess temporal drift, a cohort of 23 scans submitted during the initial portion of the CATT study was longitudinally followed with serial reproducibility analysis.

INTERVENTION—CATT readers performed standardized grading on OCT images. A reader team, composed of two independent readers and a Senior Reader, evaluated each scan. Grading included the CATT OCT endpoints of total thickness at the foveal center point and intraretinal fluid, subretinal fluid, and sub-retinal pigment epithelium (RPE) fluid. Independent reading teams masked to the results of initial grading reevaluated scans to determine reproducibility of qualitative grading and measurements.

MAIN OUTCOME MEASURES—Categorical grading agreement was reported using percent agreement and kappa statistic and measurement agreement was reported using intraclass correlations and paired differences.

RESULTS—Reading Center teams reproducibly graded intraretinal fluid (percent agreement = 73%, kappa = 0.48, 95% confidence interval or CI 0.38 to 0.58), subretinal fluid (percent agreement = 90%, kappa = 0.80, 95% CI 0.73 to 0.87), and sub-RPE fluid (percent agreement 88%, kappa = 0.75, 95% CI 0.67 to 0.83). For independent Reading Center team measurements of

© 2012 American Academy of Ophthalmology, Inc. Published by Elsevier Inc. All rights reserved.

Corresponding Author: Glenn J. Jaffe, MD, Duke Eye Center, Director of Duke Reading Center, DUMC Box 3802, Durham, NC 27710, jaffe001@mc.duke.edu (may be published), Phone: 919-684-4458 Fax: 919-681-6474.
Reprint Address: Glenn J. Jaffe, MD, DUMC Box 3802, Durham, NC 27710

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.

Presented at Association for Research in Vision and Ophthalmology (ARVO) May 2012, Ft. Lauderdale, FL.

Conflict of Interest: No conflicting relationship exists for any author.

This article contains online-only material. The following should appear online-only: Table 1, Table 2, Table 3, Table 4, and Figure 5.

total thickness at the foveal center point, the intraclass correlation was 0.99 (95% CI 0.99 to 0.99, and the mean paired difference between Reading Center teams was -4 micrometers (95% limits of agreement -55 to 47 micrometers). There was no qualitative or quantitative grading drift.

CONCLUSIONS—The standardized protocols used to evaluate OCT scans from the CATT study were reproducible. The methods used are suitable to monitor OCT imaging data from a large neovascular age related macular degeneration interventional multicenter study.

INTRODUCTION

The Comparison of Age-Related Macular Degeneration Treatments Trials (CATT) is a prospective, randomized, multi-center, clinical trial that compares the relative safety and efficacy of intravitreal bevacizumab to intravitreal ranibizumab as interventions for neovascular Age-Related Macular Degeneration (NVAMD)¹. This trial also examines the relative efficacy of different dosing schedules of each agent. Various imaging modalities including optical coherence tomography (OCT, Carl Zeiss Meditec, Dublin, California) were used to monitor CATT study patient response to therapy.

OCT provides a non-invasive way to obtain cross sectional images of the retina. Anatomic changes associated with NVAMD such as intraretinal fluid, subretinal fluid, hyper-reflective material under the retina, and pigment epithelial detachment (PED) can be readily visualized on OCT²⁻⁴. These pathologic changes⁵ as well as the efficacy of various treatments⁶⁻⁹ can be followed longitudinally on OCT. Furthermore, OCT-facilitated determination of presence or absence of macular fluid associated with choroidal neovascularization has also been used to rationally direct intravitreal pharmacologic therapy¹⁰⁻¹³.

In the CATT, macular fluid, defined as one or more of the following, intraretinal fluid (IRF), subretinal fluid (SRF), and/or sub-retinal pigment epithelium (RPE) fluid, was an eligibility prerequisite, re-treatment criteria, and secondary study endpoint. Accordingly, it is important to accurately and reproducibly identify macular fluid to ensure appropriate study enrollment and treatment, and to correctly interpret study results.

To evaluate CATT OCT images, we adopted a novel team based grading approach; a pool of CATT Readers was chosen and two readers selected from this pool independently graded each scan. Any discrepancies between the two readers were arbitrated by a Senior Reader. Herein, we report the reproducibility of the CATT OCT grading protocol, and whether the grading changed over time.

MATERIALS AND METHODS

Approval for this study was obtained from the Duke Institutional Review Board. All experimental procedures adhered to the tenets of the Declaration of Helsinki, and all participants engaged in an informed consent process and signed a written consent document prior to enrollment in the CATT (ClinicalTrials.gov Identifier: NCT00593450). For the CATT, the qualitative OCT endpoint was the presence of macular fluid and the quantitative endpoint was thickness at the foveal center. A description of OCT acquisition procedures, site technician and reader certification, and grading methodology can be found in Chapter 18 of the CATT manual of Procedures (<http://www.med.upenn.edu/cpob/studies/CATT.shtml>. Accessed May 26, 2012).

Reader Certification

Certified readers reviewed all scans. To become certified, readers were required to review an OCT grading manual, complete a training curriculum, pass an OCT reader knowledge

assessment test, and be closely supervised by a Senior Reader until grading was determined to be accurate. When the CATT was initiated, a pool of 2 Senior Readers (readers who have fulfilled all of the training requirements of a reader, and, have additionally completed additional pre-specified advanced training activities) and 3 readers were designated as CATT Readers. As the study scan volume increased, these numbers were expanded to include a pool of 4 Senior Readers and 8 readers. At any one time 3 to 8 readers and Senior Readers concurrently analyzed study scans.

OCT Scan Acquisition

All study scans were acquired by CATT-certified OCT technicians using Stratus OCT machines. To become CATT-certified, a technician successfully completed a knowledge assessment test and received image acquisition training that emphasized appropriate focus, scan saturation, line length, and line placement. The technician submitted 16 certification scans to Reading Center imaging specialists, who evaluated the scans to verify that the scans were of high quality and were obtained according to the study scan protocol. Certification was awarded once the technician had successfully completed these requirements. An automated e-mail feedback system reported scan quality, placement, and individually identified scans of concern to OCT technicians for all scans submitted to the Reading Center during the CATT.

Prior to OCT scan submission to the Reading Center, all patient identifying data were removed in compliance with Health Insurance Portability and Accountability Act (HIPAA) guidelines. All eyes were imaged with both the fast macular thickness (FMTM) and the macular thickness map (MTM) scan protocols. Less than 1% of scans submitted to the Reading Center did not adhere to this submission protocol.

OCT Scan Grading

Each of twelve (six from FMTM and six from MTM) radial line images was assessed during grading. All OCT scans were analyzed for the presence of the following parameters: vitreomacular attachment (VMA), epiretinal membrane (ERM), intraretinal fluid (IRF), subretinal fluid (SRF), subretinal hyperreflective material (SHRM), and RPE elevation (RPEE) as depicted in Figure 1a – f and Figure 2a – f. Vitreomacular attachment was defined as vitreous attachment and focal separation from the inner retina within a 3 mm diameter centered at the middle of the fovea. Standardized reference images were compiled to illustrate examples of each morphological feature, and were made available to all CATT readers.

For each morphological feature evaluated, one of the following grades was assigned: feature present, feature absent, not interpretable (due to incorrect scan placement or poor scan saturation), or absent scan. OCT morphological features graded as present ranged from subtle to obvious. Examples of obvious and subtle morphological findings are shown in Figure 1a – f and Figure 2a – f.

If a particular OCT morphologic feature was graded as present on a scan, a reader was always required to further subcategorize the feature. For example, if ERM or VMA was graded present, a reader recorded the presence of any associated deformation of the central 1 mm of the retina (Figure 3a – b). If RPEE was present, a reader recorded whether sub-RPE fluid was present. If macular fluid (one or more of IRF, SRF, or sub-RPE fluid) was graded present, a reader determined if that specific type of macular fluid was present anywhere within the central 1 mm of the OCT scan, and also if any macular fluid was present at the foveal center point (Figure 3c – h). Finally, if SHRM was graded present, a reader

determined if SHRM was present anywhere within the central 1 mm of the retina (Figure 3i).

After morphological grading was completed, morphometric analysis was performed on each scan. Quantitative values for morphometric variables were preferentially recorded from the 6 radial line images produced by the MTM protocol, though if these were of not acceptable quality, individual images from the FMTM protocol could be substituted. The largest horizontal and vertical dimensions for RPEE were measured from each of the 6 radial line scans, and the maximum value on a single radial scan for both dimensions was reported. We defined RPEE height (vertical dimension) from Bruch's membrane to the basal RPE surface of the RPE and RPEE width (horizontal dimension) from the point where the RPE started to separate from the choroid and become elevated to the point where the RPE was flat against Bruch's membrane and was no longer elevated.

For each radial line scan evaluated, thickness (vertical dimension) at the foveal center point was reported for each of the following: retina, SRF, and choroidal neovascularization – pigment epithelial detachment (CNV-PED) complex. CNV-PED complex thickness was defined as the sum of RPE thickness, RPEE thickness, and SHRM thickness, as individual borders of these features were difficult to consistently delineate with accuracy (Figure 4). The sum of retinal thickness, SRF thickness, and CNV-PED complex thickness at the foveal center point defined the CATT quantitative OCT endpoint of total thickness at the foveal center point.

Vertical dimension measurements of retinal, SRF, and CNV-PED complex thickness were performed on all 6 radial line scans until April 2009. From April 2009 onwards, measurements of vertical dimension of the retina, SRF, and CNV-PED complex thickness were performed on all 6 radial line scans for study visits at week 000, 004, 008, 012, 024, and 052. The mean thickness measurements averaged from scans 1 and 4 were determined by the CATT Coordinating Center to be approximately equal to mean thickness measurements derived from the average of 6 radial line scans (data not shown). Thus to increase grading efficiency, and minimize unnecessary measurements, for the remaining year one CATT study visits submitted after April 2009, vertical dimensions were measured for retinal, SRF, and CNV-PED complex thickness on radial line scans 1 and 4 alone. Thickness measurements were performed manually on a standardized monitor at defined image size with a ruler and then converted to micrometers at the CATT Coordinating Center.

Team Based Grading

Two masked readers individually graded all OCT scans in parallel. An independent data transcriptionist identified discrepant values between the paired readers. Morphometric data was considered discrepant if the vertical measurement differed by more than 65 micrometers or the horizontal measurement differed by more than 220 micrometers. The Director of Grading and Senior Readers established these values for horizontal and vertical measurement discrepancies after analysis of aggregated Stratus OCT grading data from a prior interventional study on eyes with exudative AMD. All graded scan pairs with discrepant data were then presented to a Senior Reader for arbitration. During the arbitration process, a Senior Reader reconciled all discrepancies between the initial reader pair. Any concordant reader grades that were deemed inaccurate by the Senior Reader were likewise corrected. Senior Readers additionally reviewed all OCT scans for the presence of any macular fluid since this fluid was a study endpoint. Any finding or value that remained controversial after arbitration was forwarded to the Director of Grading for final decision.

Team Agreement Analysis

Grading reproducibility between several different pairs of Reading Center teams was analyzed on scans uploaded to the Reading Center between July 2009 and February 2010. From a subset of 274 scans randomly selected by computer from a comprehensive archive, 270 were available for reproducibility analysis. A pair of readers other than those that had performed the initial review, and a Senior Reader that had not performed original arbitration performed the reproducibility grading. All new readers were masked to the results of the first reading team. The values obtained by the second reading team were then compared to those obtained by the first reading team (Figure 5, available at <http://aojournal.org>). Of note, any morphological feature graded not interpretable by one reading team and graded either present or absent by another team was recorded as disagreement.

To test for grading drift over time, a subset of 23 scans uploaded during the initial portion of the CATT study underwent serial inter team agreement analysis. These reproducibility studies were performed at approximately 4 – 6 month intervals over the study duration.

Quantitative Intraretinal Fluid Analysis

From the 270 scans that underwent reproducibility analysis, both Reading Center teams agreed that IRF was present on 108 scans, and only a single Reading Center team reported IRF on 70 scans. To determine whether the single largest intraretinal cystoid hyporeflective cross sectional area differed between these two groups we performed comparative analysis of cross sectional area on 35 scans randomly selected from each group. All 6 images from the fast macular thickness map (FMTM) protocol and 6 images from the macular thickness map (MTM) protocol were reviewed to determine the largest horizontal and vertical dimensions from a single radial line image. Stratus software-based calipers were used to quantify the maximal horizontal and vertical dimensions of the single largest cross sectional area of IRF for a specific scan. Cross sectional area of single largest IRF was approximated as an ellipse using the following formula: $\text{area} = \pi \times (\text{Horizontal Dimension}/2) \times (\text{Vertical Dimension}/2)$. The sample size was calculated based on IRF area.

Quality Control

Several measures facilitated consistent analysis. First, all OCT scans were obtained by CATT certified OCT technicians from OCT machines using standardized software packages (version 4.0 or greater). Next, all data entry by transcriptionists into a centralized database was verified via an independent data entry team. Finally, ongoing monthly meetings ensured adherence to study grading protocols, addressed general discrepancies, and allowed for consensus opinion regarding controversial scans. It is worthwhile to note that only one scan of the 270 that underwent reproducibility analysis was discussed at a monthly meeting within 3 months of the actual reproducibility exercise.

Statistical Analysis

For categorical measures, the percent agreement (grading concordance between Reading Center teams) was computed to determine agreement. Percent agreement was computed as the number of concordant grading pairs divided by the total number of grading pairs multiplied by 100. Additionally, Kappa statistics and respective 95% confidence intervals were reported using the guidelines proposed by Landis and Koch^{14, 15}: greater than 0.80 = near perfect agreement, 0.61 – 0.80 = substantial agreement, 0.41 – 0.60 good agreement, and 0.21 – 0.40 fair agreement.

For assessing reproducibility of continuous measures, paired differences were computed. The mean (standard deviation) of paired difference and 95% limits of agreement was calculated. The significance of paired differences was assessed using the Wilcoxon signed

rank test of median difference equal to zero and intraclass correlations were used to summarize the agreement of continuous measures.

A Wilcoxon rank sum test of difference in medians was used to compare the difference in intraretinal fluid area for eyes with concordant intraretinal fluid grades to those with discordant grades. All analyses were performed with SAS 9.2 software (SAS, Cary, North Carolina).

RESULTS

Reading Center Team Agreement

Table 1 (available at <http://aaojournal.org>) summarizes grading agreement between Reading Center teams for evaluation of all OCT morphologic features in 270 OCT scans. Percent agreement between grading teams for macular fluid was 84%. Percent agreement for IRF, SRF, and sub-RPE fluid was 73%, 90%, and 88% respectively. Independent Reading Center teams demonstrated good or better levels of agreement, based on kappa statistics, for grading of morphological features. For IRF, SRF, and sub-RPE fluid, kappa statistics were 0.48, 0.80 and 0.75 respectively. The kappa statistic for macular fluid was 0.55.

Table 2 (available at <http://aaojournal.org>) details the agreement between Reading Center teams for all OCT quantitative measurements. For mean total thickness at the foveal center point, the intraclass correlation between Reading Center teams was 0.99 (95% confidence interval or CI 0.99 to 0.99). For mean retinal thickness at the foveal center point, mean subretinal fluid thickness at the foveal center point, and mean CNV-PED complex (RPE +RPEE+SHRM) thickness at the foveal center point, the intraclass correlations between Reading Center teams were 0.93, 0.90, and 0.98 respectively. For total thickness at the foveal center point, the mean paired difference between Reading Center teams was -4 micrometers (95% limits of agreement -55 to 47 micrometers). For mean retinal, subretinal fluid, and CNV-PED complex (RPE+RPEE+SHRM) thickness at the foveal center point, paired differences (95% limits of agreement in micrometers) between Reading Center teams were -3 (-62 to 56), 0.6 (-27 to 28), and -2 (-61 to 57) micrometers respectively. The mean paired differences between Reading Center teams for all OCT measurements are shown in Table 2 (available at <http://aaojournal.org>).

Analysis of Temporal Drift Grading

Serial grading of a cohort of scans demonstrated comparable levels of inter team agreement over time (Table 3, available at <http://aaojournal.org>). For macular fluid, percent agreement ranged from 78% - 83%. For IRF, SRF, and sub-RPE fluid, percent agreement was 57 - 70%, 83 - 100%, and 78 - 91% and respectively.

For mean total thickness at the foveal center point, intraclass correlations between Reading Center teams over time were 0.97 at all 3 time points. For mean retinal thickness at the foveal center point, mean subretinal fluid thickness at the foveal center point, and mean CNV-PED complex (RPE+RPEE+SHRM) thickness, the intraclass correlations between Reading Center teams over time ranged between 0.95 - 0.97, 0.98 - 1.00, and 0.97 - 0.98 respectively. The agreement for each morphometric feature undergoing longitudinal analysis is shown in Table 4 (available at <http://aaojournal.org>). For mean total thickness at the foveal center point, the mean paired differences between Reading Center teams over time ranged between -10 to -3 micrometers. For mean retinal thickness at the foveal center point, mean subretinal fluid thickness at the foveal center point, and mean CNV-PED complex (RPE+RPEE+SHRM) thickness at the foveal center point, the mean paired differences between Reading Center teams over time ranged between -1 to -0.2, 0 to 0, and -9 to -1 micrometers respectively. The mean paired measurement differences for all OCT

measurements undergoing longitudinal analysis can be found in Table 4 (available at <http://aojournal.org>).

Quantitative Intraretinal Fluid Analysis

Median single largest intraretinal cystoid hyporeflective cross sectional area (median = $11.7 \times 10^{-3} \text{mm}^2$, range $1.9 - 135.0 \times 10^{-3} \text{mm}^2$) on 35 randomly sampled scans where Reading Center teams agreed on IRF presence was larger ($p = 0.001$) when compared to median single largest intraretinal cystoid hyporeflective cross sectional area (median = $5.5 \times 10^{-3} \text{mm}^2$, range $1.3 - 570.8 \times 10^{-3} \text{mm}^2$) on 35 randomly sampled scans where only one Reading Center team graded IRF as present. Representative images depicting median single largest intraretinal cystoid hyporeflective cross sectional area for scans where both Reading Center team agreed on presence of IRF and where only one Reading Center team reported fluid are shown in Figure 6.

DISCUSSION

In the present study, we have shown that well trained reader teams in a Reading Center setting can reproducibly grade OCT qualitative and quantitative features in a large multicenter randomized interventional neovascular AMD treatment trial. Of the CATT OCT endpoint macular fluid variables, agreement was best for subretinal and sub-RPE fluid. Reproducibility was generally excellent for quantitative parameters. We believe that the reproducible results that we obtained resulted from rigorous reader certification requirements, collectively understood definitions of morphological characteristics, and consistently applied quantitative measurement protocols.

Previously, we have shown that OCT images generated from 132 eyes in an interventional neovascular AMD trial were reproducibly interpreted in a Reading Center setting¹⁶. In the present study, we observed 73% and 90% team grading agreement for IRF and SRF respectively, comparable to the 84 – 85% and 90 – 91% inter-reader agreement for IRF and SRF respectively that we reported previously. For total thickness measurement at the foveal center in the current work we noted a median paired difference of 0 micrometers between teams, which was less than the 21 – 64 micrometer range of inter-reader median measurement differences reported in the previous study. These modest disparities may be due to differences in the trial enrollment criteria, OCT scan acquisition protocol, and grading methodology.

Our Reading Center has established a team based grading approach that includes arbitration by a Senior Reader to maximize grading consistency during the study. This process also allows a Senior Reader to review a higher volume of scans, and to establish a close feedback loop with newer readers to enhance grading consistency. Prior series detailing OCT grading protocols have utilized individual readers^{17, 18}, and paired readers in parallel^{16, 19}, while other large clinical trials employing OCT grading by a reading center have not published detailed grading protocols^{20–22}.

The “double grading” protocol for baseline fundus photographs utilized in the ETDRS study most resembles our team based OCT scan grading protocol. During the ETDRS study, only baseline color fundus photos underwent review by a pair of independent readers. One step of disagreement (out of 3 possible steps in the ETDRS fundus photo grading scale) was averaged together, and two steps or more of disagreement was returned to the initial graders for repeat evaluation. A masked ETDRS senior grader resolved any persistent disagreements. For subsequent study visits, a single reader alone evaluated follow up fundus photos, and grading was monitored using “haphazardly selected reading lists” of 10 eyes each²³. Our image grading protocol differed in that an independent grading team evaluated

both baseline and follow up images, Senior Readers arbitrated all grading inconsistencies, and reproducibility studies were systematically performed on Reading Center teams. Though ETDRS “double grading” has similarities to our team based grading protocol, our evaluation methods more stringently address grading discrepancies and reproducibility.

Reading Center grading was reproducible on morphological features. Agreement was highest for subretinal fluid and less for intraretinal fluid and epiretinal membrane. Cystoid hyporeflective areas within the retina on OCT represent intraretinal fluid from NVAMD^{24, 25}. However, a variety of factors may compromise intraretinal fluid identification. There may be increased hyporeflective pixels within the retina, which in hyporeflective layers of the retina may have the appearance of small cystoid changes when none are actually present on scans with low signal intensity due to media opacity, low signal strength, or other factors. We have termed this finding a “pixel void”. Even the normal foveal center often appears slightly hyporeflective on OCT and can mimic very subtle intraretinal fluid, especially when coupled with decreased scan signal intensity. Finally, underlying active choroidal neovascular membranes may result in subretinal fluid at the CNV-retinal interface making it difficult to discriminate intraretinal fluid from subretinal fluid.

We found that the single largest cystoid hyporeflective area was smaller when only one Reading Center team reported fluid. It is not surprising that smaller true cystoid spaces are more challenging to grade consistently. These smaller areas of fluid are more difficult to differentiate from pixel voids, than those with a larger cross-sectional area.

Epiretinal membranes can be difficult to visualize on Stratus OCT, especially when tractional changes are not visualized at the inner retina. In addition, a jagged, discontinuous inner retinal boundary that mimics an epiretinal membrane can be seen when OCT image saturation is decreased. Hallmarks of epiretinal membrane such as focal points of attachment, optical reflectivity difference, and visible tufts or edges²⁶ may not be visible on Stratus OCT during grading. One group reported a 30% increase in epiretinal membrane detection rate when using ultrahigh resolution spectral domain OCT (SD-OCT) compared to Stratus OCT²⁷.

Reading Center teams demonstrated high levels of quantitative grading agreement. For all thickness measurements at the foveal center point, we observed relatively small mean paired thickness measurement differences less than 5 micrometers and high intraclass correlations between 0.90 – 0.99. For the trial endpoint total thickness at the foveal center point, the mean (\pm standard deviation or SD) of paired difference was 3.9 ± 25.7 micrometers ($p = 0.025$). Though this difference was statistically significant, a Reading Teams measurement difference of less than 4 micrometers is likely not clinically significant. These minimal differences and high levels of measurement agreement are especially notable in light of Bruch’s membrane obscuration by overlying CNV or disruption of the RPE layer by CNV. These pathological changes common to NVAMD can make accurate segmentation of the outer retina more difficult. To minimize these segmentation difficulties, our protocol aggregated RPE thickness, any RPE elevation, and subretinal hyperreflective material thickness as a single measurement termed “CNV-PED complex” thickness.

Reading Center teams also demonstrated excellent agreement when measuring maximal RPE elevation height (intraclass correlation = 0.97) and lower agreement when grading maximal RPE elevation width (intraclass correlation = 0.81). The heterogeneous changes induced by CNV in the subretinal space as visualized on OCT may partly account for the reduced reproducibility in grading RPE elevation width. For example, within an area of RPE elevation, CNV-mediated RPE fragmentation can make it difficult to consistently identify

the exact separation point of the RPE from Bruch's membrane. Additionally, overlying SHRM can sometimes obscure the borders of underlying RPE elevation. Finally, in scans with multiple adjacent RPE elevations, it can be challenging to confirm if a single RPE elevation is discrete or contiguous with adjacent RPE elevations due to difficulty visualizing each potential point of RPE attachment to Bruch's membrane.

We evaluated reader agreement over time in a cohort of subjects followed from the initiation of CATT to monitor temporal grading drift. No obvious temporal drift was identified. We hypothesize that ongoing reader training and feedback during the study helped to minimize variations in Reader grading over time.

There are limitations to this study. The data were derived from a single Reading Center. Accordingly, reader reproducibility reported herein may not be readily generalized to other reading centers. Nonetheless, we believe that readers in other settings could adopt our team-based approach, with ongoing reader training and feedback, and standardized grading protocols to produce reproducible grading data. In fact, prior work demonstrated generally high levels of OCT grading agreement between independently trained reader pairs at two different reading centers¹⁹. Next, a Senior Reader reviewed all scans analyzed by primary readers for intraretinal, subretinal, and sub-RPE fluid, key morphological variables in the CATT. However, for other morphological variables, if the grade assigned by the two primary readers was not discrepant, the Senior Reader did not necessarily review the scans. Accordingly, it is conceivable that if a variable was ascribed an identical inaccurate value by both primary readers, the Senior Reader might not correct the inaccuracy. However, we believe that these instances are likely rare, and, for several reasons would have minimal impact on the study results. First, a reader was not consistently matched with a particular second reader. Though two individuals may make similar grading errors, the likelihood of several readers all making an identical error for the same grading variable is small. Next, independent Reading Center teams showed high levels of agreement with one another. Discounting widespread and systematic biases across the entire Reading Center, the chances of four to six independent readers obtaining identical erroneous values for a particular finding is low. Finally, Senior Readers corrected erroneous values consistently reported by a reader pair if these values were determined to be inaccurate during arbitration. These scans were then returned to the reader pair for mandatory review, to maintain grading consistency across readers.

During categorical grading analysis of all OCT morphological features we reported both percent agreement and kappa statistic in consideration of the innate limitations of this second analysis method. In particular, case distribution could result in high percent agreement but low values for kappa statistic. In the event that cases are very common or very rare, kappa statistic can differ widely from percent agreement^{28, 29}. This phenomenon was apparent in the current study for less commonly observed morphological feature such as vitreomacular adhesion (94% agreement, kappa = 0.74) and epiretinal membrane (95% agreement, kappa 0.53). The disparity between percent agreement and kappa statistic was more pronounced for the even less frequently observed grading variables vitreomacular adhesion with foveal deformation (82% agreement, kappa = 0.49) and epiretinal membrane with foveal deformation (90% agreement, kappa = 0.46).

Future investigations will capitalize upon the numerous advantages offered by spectral domain OCT technology (SD-OCT). Compared to conventional time domain OCT (TD-OCT), such as Stratus OCT used for this study, SD-OCT offers increased image resolution, improved registration, and faster data acquisition resulting in decreased motion artifact^{30, 31}. These advantages may result in increased detection of important retinal features such as IRF, SRF, and sub-RPE fluid^{32,33}. If so, reader reproducibility may have been even higher than

that reported in the present study. A SD-OCT sub-study has been initiated in CATT, and definitive answers to questions regarding reader reproducibility with spectral domain OCT when compared with time domain OCT will be forthcoming when the sub-study has been completed.

Since clinical studies for retinal diseases increasingly incorporate OCT to better understand treatment effect, reproducible analysis of imaging data is crucial to understand the efficacy of an intervention and to consistently evaluate an individual's response to therapy. This study demonstrates that Reading Center teams can reproducibly grade OCT images to facilitate monitoring of therapeutic effect in a large, prospective, multi-center, interventional treatment trial for NVAMD. A standardized training, grading, and feedback protocol can employ readers with differing levels of experience and obtain consistent results, while maintaining quality over time.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Financial Support: NIH 5U10EYO17825. This work was also supported in part by the Heed Foundation.

REFERENCES

1. Martin DF, Maguire MG, Ying GS, et al. Ranibizumab and bevacizumab for neovascular age-related macular degeneration. *N Engl J Med*. 2011; 364:1897–1908. [PubMed: 21526923]
2. Hee MR, Bauman CR, Puliafito CA, et al. Optical coherence tomography of age-related macular degeneration and choroidal neovascularization. *Ophthalmology*. 1996; 103:1260–1270. [PubMed: 8764797]
3. Jaffe GJ, Caprioli J. Optical coherence tomography to detect and manage retinal disease and glaucoma. *Am J Ophthalmol*. 2004; 137:156–169. [PubMed: 14700659]
4. Ting TD, Oh M, Cox TA, et al. Decreased visual acuity associated with cystoid macular edema in neovascular age-related macular degeneration. *Arch Ophthalmol*. 2002; 120:731–737. [PubMed: 12049577]
5. Hee MR, Puliafito CA, Wong C, et al. Quantitative assessment of macular edema with optical coherence tomography. *Arch Ophthalmol*. 1995; 113:1019–1029. [PubMed: 7639652]
6. Rogers AH, Martidis A, Greenberg PB, Puliafito CA. Optical coherence tomography findings following photodynamic therapy of choroidal neovascularization. *Am J Ophthalmol*. 2002; 134:566–576. [PubMed: 12383814]
7. Kaiser PK, Blodi BA, Shapiro H, Acharya NR. Angiographic and optical coherence tomographic results of the MARINA study of ranibizumab in neovascular age-related macular degeneration. *Ophthalmology*. 2007; 114:1868–1875. [PubMed: 17628683]
8. Avery RL, Pieramici DJ, Rabena MD, et al. Intravitreal bevacizumab (Avastin) for neovascular age-related macular degeneration. *Ophthalmology*. 2006; 113:363–372. e5. [PubMed: 16458968]
9. Rich RM, Rosenfeld PJ, Puliafito CA, et al. Short-term safety and efficacy of intravitreal bevacizumab (Avastin) for neovascular age-related macular degeneration. *Retina*. 2006; 26:495–511. [PubMed: 16770255]
10. Gupta OP, Shienbaum G, Patel AH, et al. A Treat and Extend Regimen Using Ranibizumab for Neovascular Age-Related Macular Degeneration Clinical and Economic Impact. *Ophthalmology*. 2010; 117:2134–2140. [PubMed: 20591490]
11. Dadgostar H, Ventura AA, Chung JY, et al. Evaluation of injection frequency and visual acuity outcomes for ranibizumab monotherapy in exudative age-related macular degeneration. *Ophthalmology*. 2009; 116:1740–1747. [PubMed: 19643484]

12. Rothenbuehler SP, Waeber D, Brinkmann CK, et al. Effects of ranibizumab in patients with subfoveal choroidal neovascularization attributable to age-related macular degeneration. *Am J Ophthalmol.* 2009; 147:831–837. [PubMed: 19217019]
13. Fung AE, Lalwani GA, Rosenfeld PJ, et al. An optical coherence tomography-guided, variable dosing regimen with intravitreal ranibizumab (Lucentis) for neovascular age-related macular degeneration. *Am J Ophthalmol.* 2007; 143:566–583. [PubMed: 17386270]
14. Koch GG, Landis JR, Freeman JL, et al. A general methodology for the analysis of experiments with repeated measurement of categorical data. *Biometrics.* 1977; 33:133–158. [PubMed: 843570]
15. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* 1977; 33:159–174. [PubMed: 843571]
16. Zhang N, Hoffmeyer GC, Young ES, et al. Optical coherence tomography reader agreement in neovascular age-related macular degeneration. *Am J Ophthalmol.* 2007; 144:37–44. [PubMed: 17509508]
17. Krebs I, Hagen S, Brannath W, et al. Repeatability and Reproducibility of Retinal Thickness Measurements by Optical Coherence Tomography in Age-Related Macular Degeneration. *Ophthalmology.* 2010; 117:1577–1584. [PubMed: 20557937]
18. Domalpally A, Blodi BA, Scott IU, et al. The Standard Care vs Corticosteroid for Retinal Vein Occlusion (SCORE) study system for evaluation of optical coherence tomograms: SCORE study report 4. *Arch Ophthalmol.* 2009; 127:1461–1467. [PubMed: 19901211]
19. Ritter M, Elledge J, Simader C, et al. Evaluation of optical coherence tomography findings in age-related macular degeneration: a reproducibility study of two independent reading centres. *Br J Ophthalmol.* 2011; 95:381–385. [PubMed: 20805123]
20. Glassman AR, Beck RW, Browning DJ, et al. Comparison of optical coherence tomography in diabetic macular edema, with and without reading center manual grading from a clinical trials perspective. *Invest Ophthalmol Vis Sci.* 2009; 50:560–566. [PubMed: 18566462]
21. Haller JA, Bandello F, Belfort R Jr, et al. Randomized, sham-controlled trial of dexamethasone intravitreal implant in patients with macular edema due to retinal vein occlusion. *Ophthalmology.* 117:1134–1146. e3. [PubMed: 20417567]
22. Brown DM, Campochiaro PA, Singh RP, et al. Ranibizumab for macular edema following central retinal vein occlusion: six-month primary end point results of a phase III study. *Ophthalmology.* 117:1124–1133. e1. [PubMed: 20381871]
23. ETDRS. Grading diabetic retinopathy from stereoscopic color fundus photographs--an extension of the modified Airlie House classification. ETDRS report number 10. Early Treatment Diabetic Retinopathy Study Research Group. *Ophthalmology.* 1991; 98:786–806. [PubMed: 2062513]
24. Ahlers C, Michels S, Elsner H, et al. Topographic angiography and optical coherence tomography: a correlation of imaging characteristics. *Eur J Ophthalmol.* 2005; 15:774–781. [PubMed: 16329065]
25. Coscas F, Coscas G, Souied E, et al. Optical coherence tomography identification of occult choroidal neovascularization in age-related macular degeneration. *Am J Ophthalmol.* 2007; 144:592–599. [PubMed: 17698019]
26. Wilkins JR, Puliafito CA, Hee MR, et al. Characterization of epiretinal membranes using optical coherence tomography. *Ophthalmology.* 1996; 103:2142–2151. [PubMed: 9003350]
27. Falkner-Radler CI, Glittenberg C, Hagen S, et al. Spectral-domain optical coherence tomography for monitoring epiretinal membrane surgery. *Ophthalmology.* 117:798–805. [PubMed: 20045567]
28. Feinstein AR, Cicchetti DV. High agreement but low kappa: I. The problems of two paradoxes. *J Clin Epidemiol.* 1990; 43:543–549. [PubMed: 2348207]
29. Crewson PE. Reader agreement studies. *AJR Am J Roentgenol.* 2005; 184:1391–1397. [PubMed: 15855085]
30. Srinivasan VJ, Wojtkowski M, Witkin AJ, et al. High-definition and 3-dimensional imaging of macular pathologies with high-speed ultrahigh-resolution optical coherence tomography. *Ophthalmology.* 2006; 113:2054 e1–2054 e14. [PubMed: 17074565]
31. Wojtkowski M, Bajraszewski T, Gorczynska I, et al. Ophthalmic imaging by spectral optical coherence tomography. *Am J Ophthalmol.* 2004; 138:412–419. [PubMed: 15364223]

32. Keane PA, Bhatti RA, Brubaker JW, et al. Comparison of clinically relevant findings from high-speed fourier-domain and conventional time-domain optical coherence tomography. *Am J Ophthalmol.* 2009; 148:242–248. e1. [PubMed: 19427620]
33. Sayanagi K, Sharma S, Yamamoto T, Kaiser PK. Comparison of spectral-domain versus time-domain optical coherence tomography in management of age-related macular degeneration with ranibizumab. *Ophthalmology.* 2009; 116:947–955. [PubMed: 19232732]

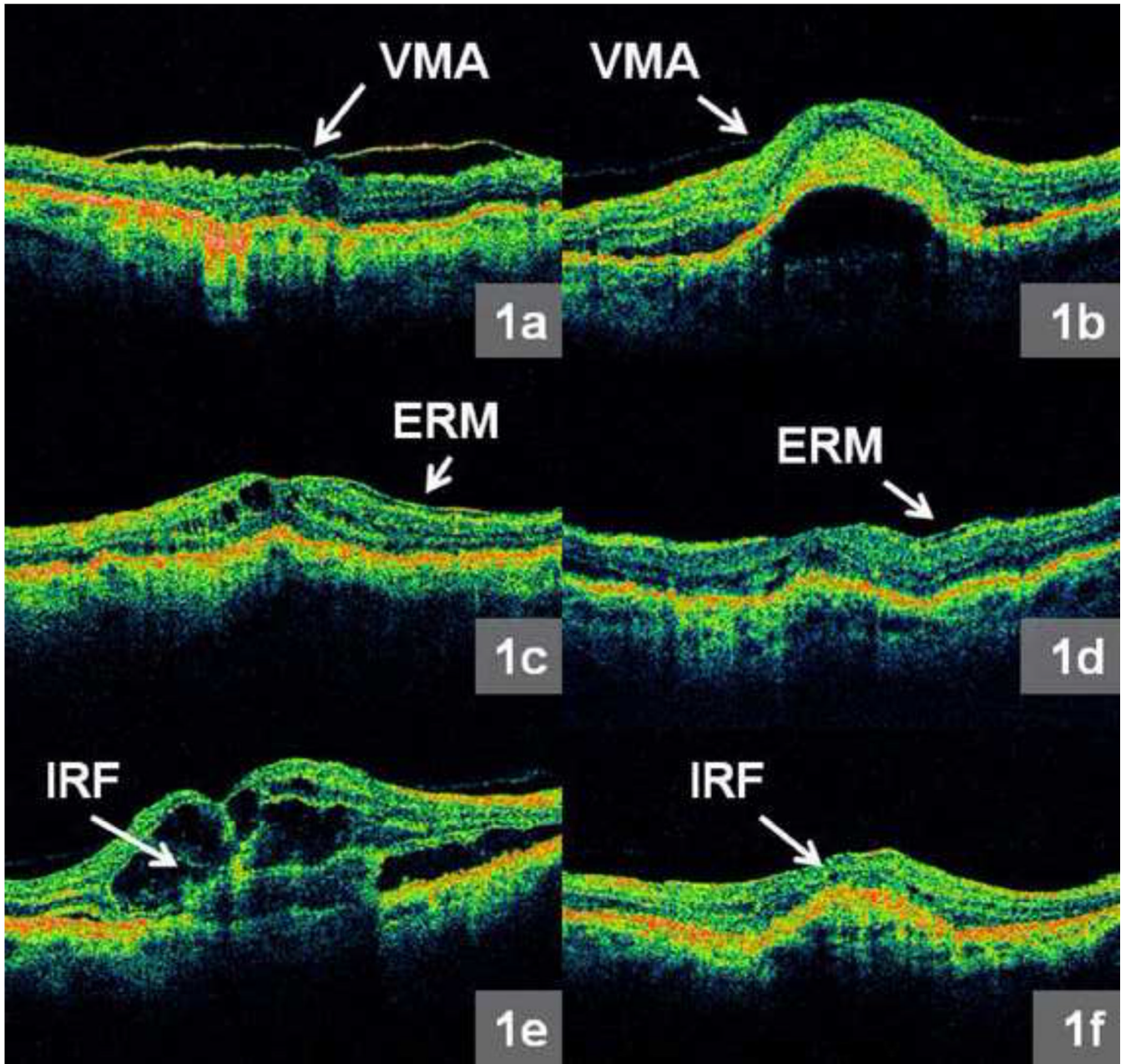


Figure 1.

a – f: Representative morphologic features from Optical Coherence Tomography (OCT) images produced by the macular thickness map (MTM) protocol: 1a. Obvious vitreomacular attachment (VMA) - vitreous attachment and focal separation from the inner retina within a 3 mm diameter horizontal region centered at the middle of the fovea, 1b. Subtle VMA, 1c. Obvious epiretinal membrane (ERM), 1d. Subtle ERM, 1e. Obvious intraretinal fluid (IRF), and 1f. Subtle IRF.

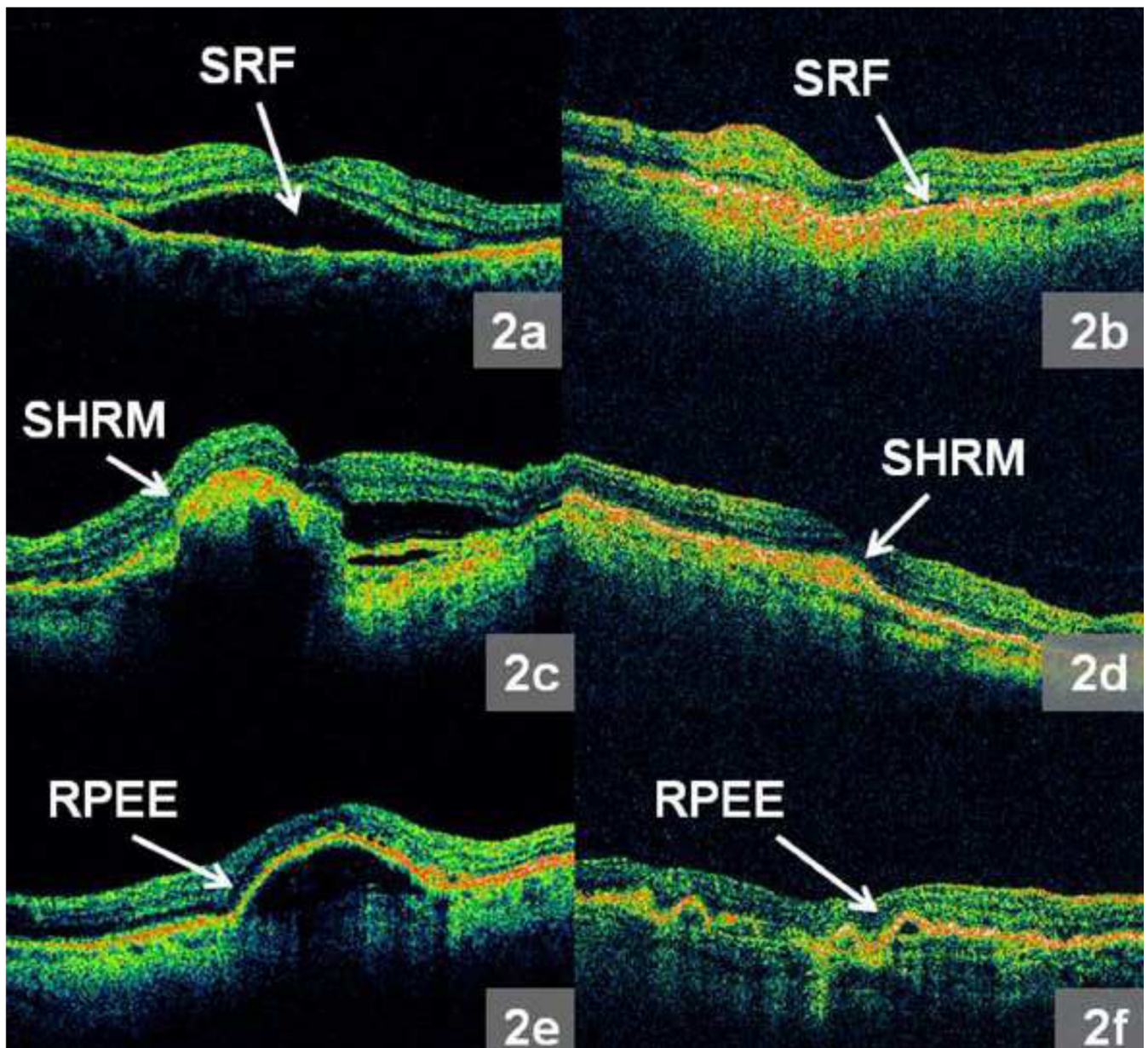


Figure 2.
a – f: Representative morphologic features from Optical Coherence Tomography (OCT) images produced by the macular thickness map (MTM) protocol: 2a. Obvious subretinal fluid (SRF), 2b. Subtle SRF, 2c. Obvious subretinal hyperreflective material (SHRM), 2d. Subtle SHRM, 2e. Obvious retinal pigment epithelium elevation (RPEE), 2f. Subtle RPEE.

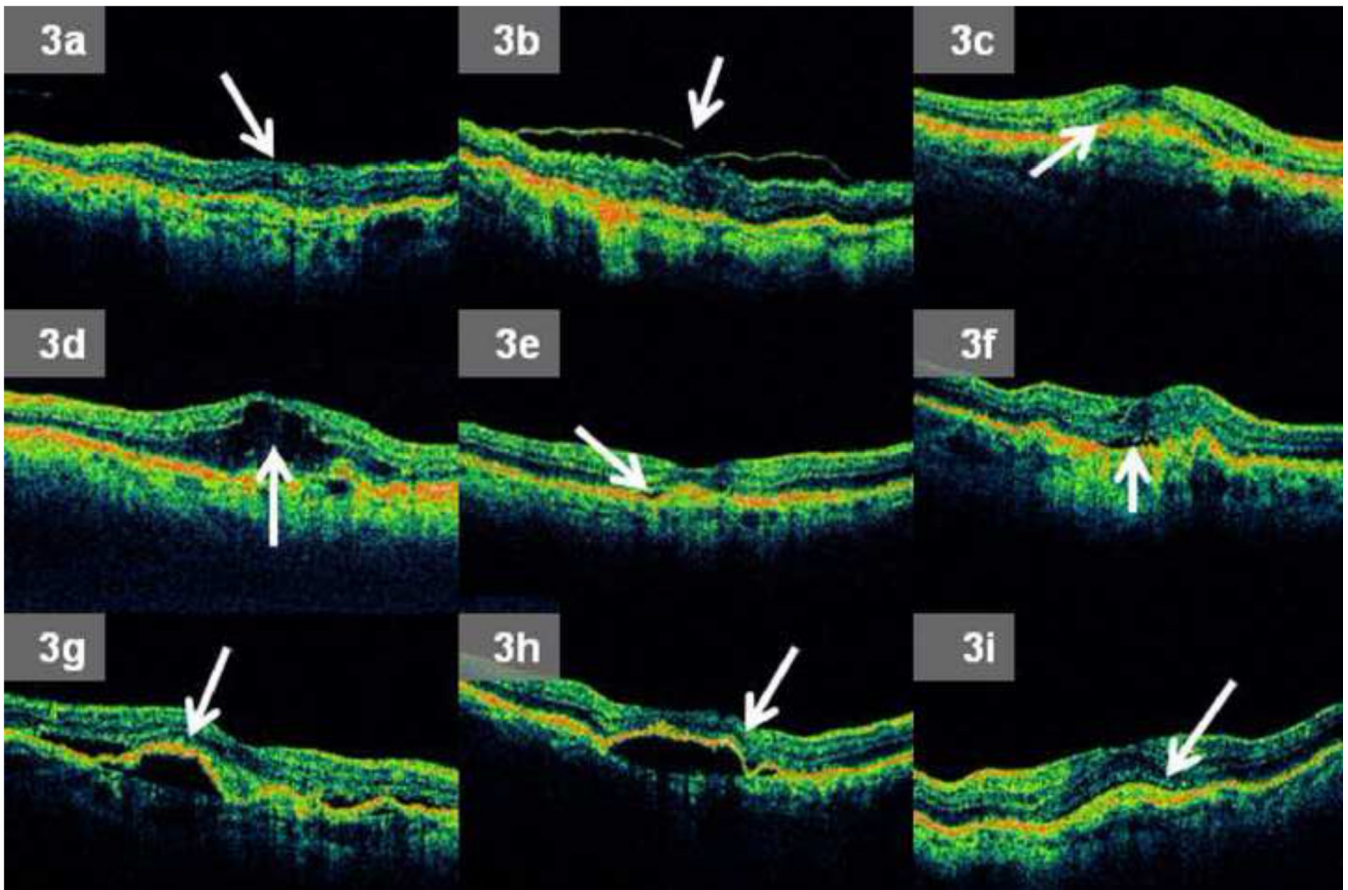


Figure 3.

a – i: Examples of grading subcategories for morphologic features noted on Optical Coherence Tomography (OCT) images produced by the macular thickness map (MTM) protocol: 3a. Epiretinal membrane (ERM) present with any deformation of the central 1 mm (horizontal dimension) of the retina, 3b. Vitreomacular attachment (VMA) present with any deformation of the central 1 mm of the retina, 3c. Any intraretinal fluid (IRF) present within central 1 mm of the retina, 3d. IRF present at the foveal center point, 3e. Any subretinal fluid (SRF) present within central 1 mm of the retina, 3f. SRF present at the foveal center point, 3g. Any sub-retinal pigment epithelium (RPE) fluid present within central 1 mm of the retina, 3h. Sub-RPE fluid present at the foveal center point, 3i. Any subretinal hyperreflective material (SHRM) present within central 1 mm of the retina.

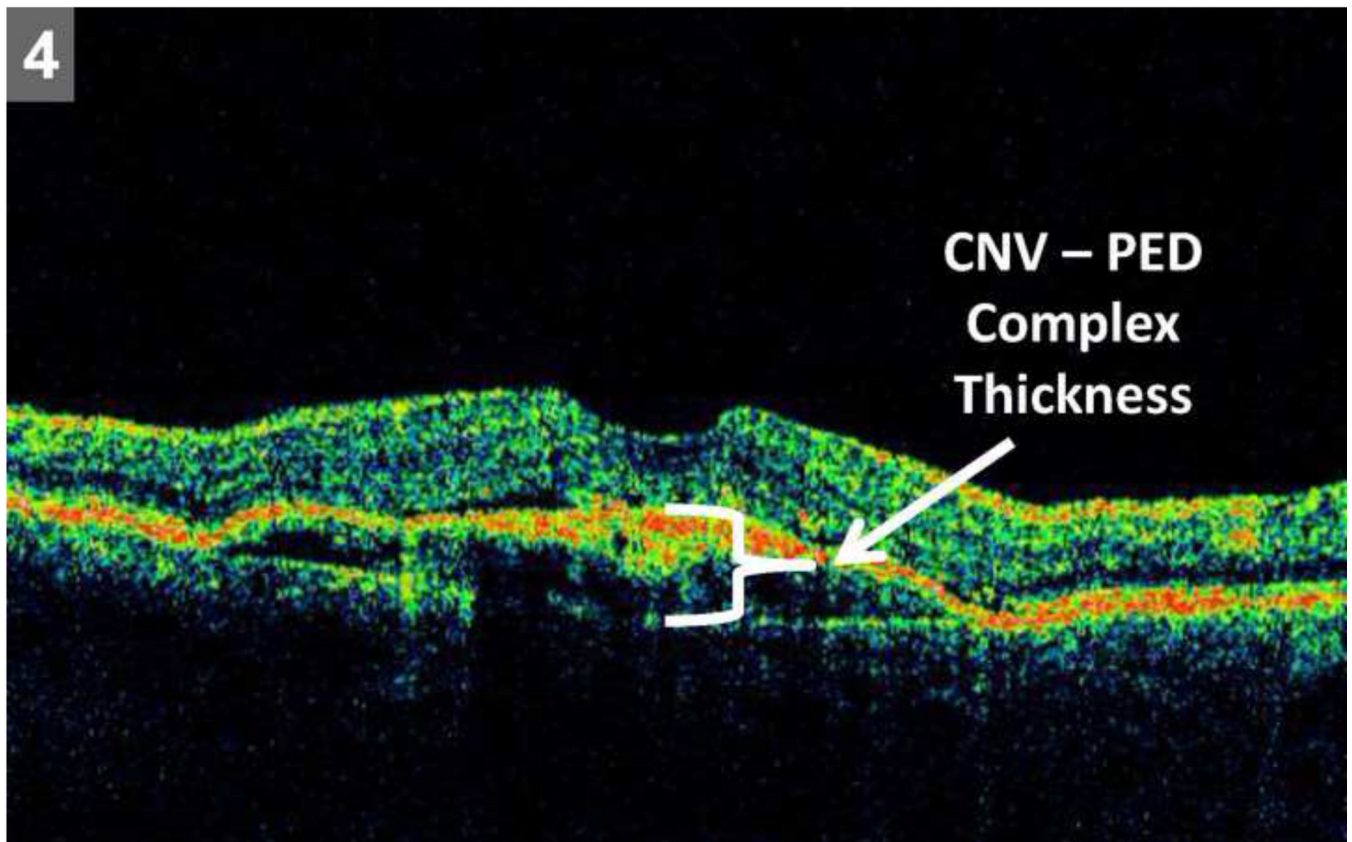


Figure 4. Representative scan demonstrating choroidal neovascular membrane – pigment epithelial detachment (CNV – PED) thickness measurement at foveal center point. The measurement was performed from the outer boundary of Bruch’s membrane to the inner boundary of the CNV.

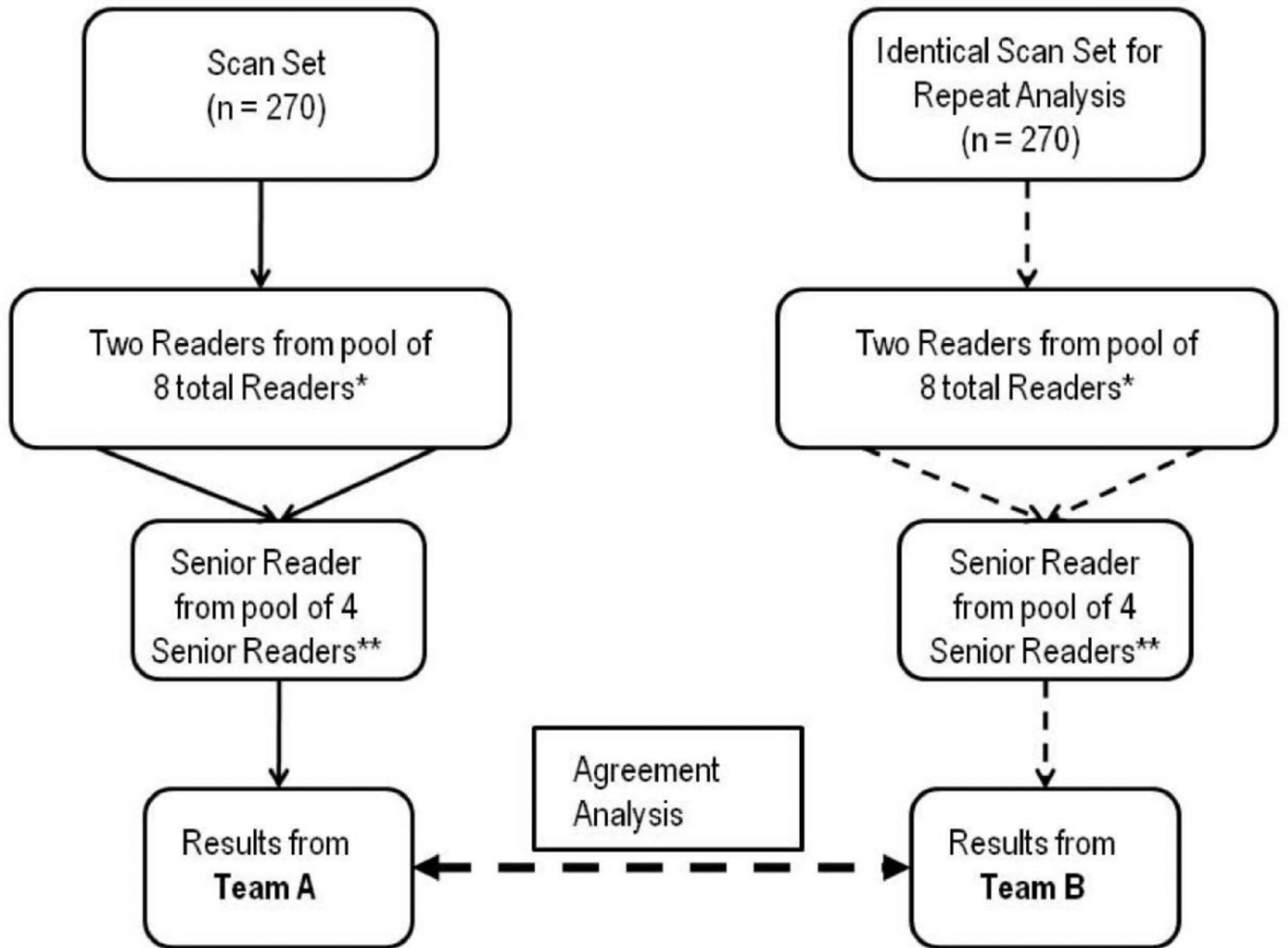


Figure 5.

OCT scan work-flow demonstrated for reproducibility analysis between hypothetical Reading Center “Team A” and “Team B” when grading the same scan. All team analyses were performed more than 5 weeks after initial reading.

solid lines = initial grading, dashed lines = reproducibility analysis. *Readers were randomly selected from pool of 8 readers, and no individual reader was part of both team A and B.

**Senior Readers were randomly selected from pool of 4 Senior Readers, and no individual Senior Reader was part of both team A and B.

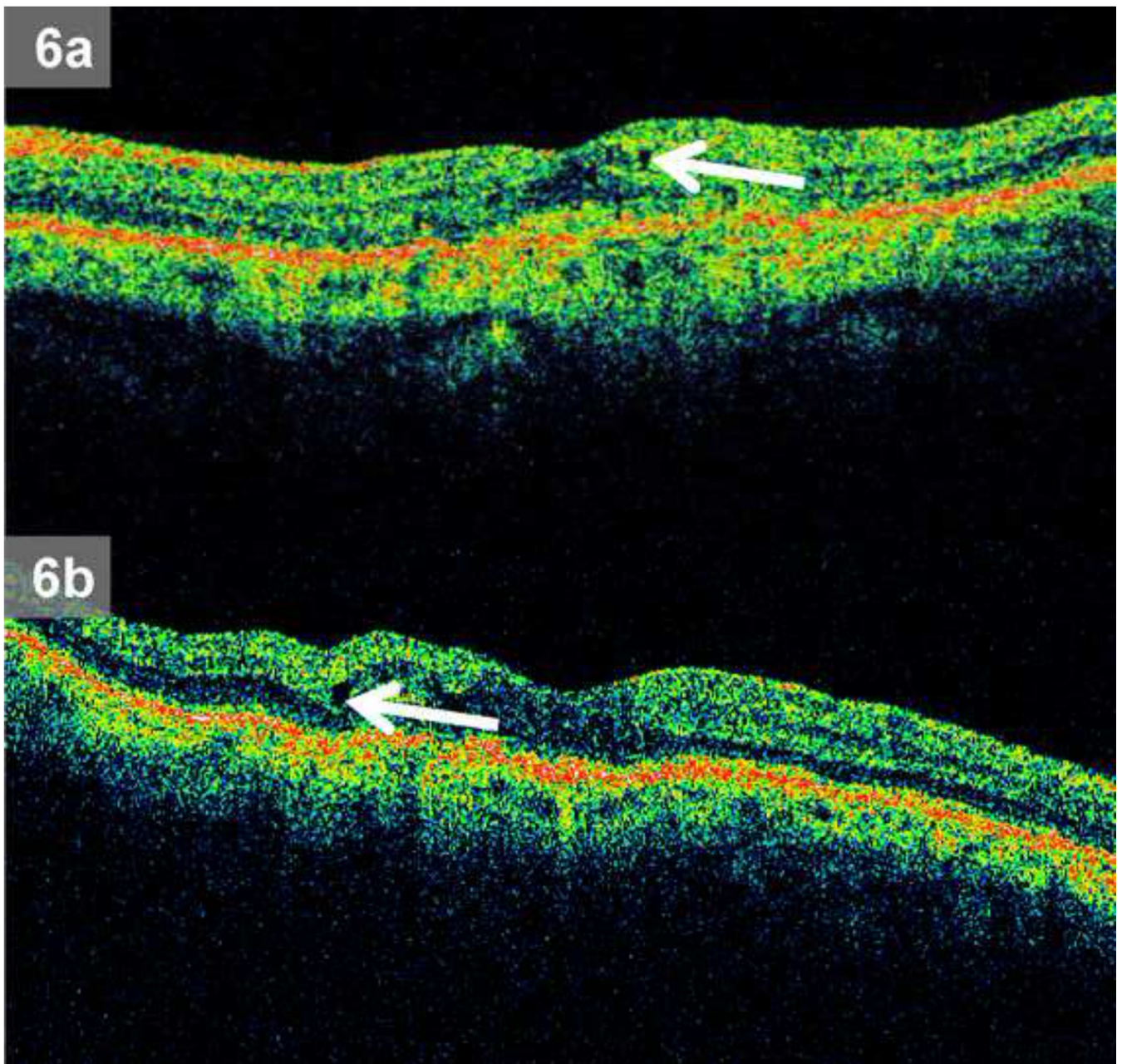


Figure 6.
a – b: Comparison of single largest intraretinal cystoid hyporeflective cross sectional area from optical coherence tomography scans: 6a. Median single largest intraretinal cystoid hyporeflective cross sectional area where only one Reading Team reported intraretinal fluid as present, 6b. Median single largest intraretinal cystoid hyporeflective cross sectional area for scans where both Reading Teams agreed on presence of intraretinal fluid.

Table 1

Reading Center Team grading agreement for qualitative OCT evaluation

Variable	Total	Percent Agreement	Kappa Statistic	Kappa Statistic Lower 95% CL	Kappa Statistic Upper 95% CL
Vitreomacular attachment	269*	94	0.74	0.63	0.86
Vitreomacular attachment with foveal deformation	269	96	0.53	0.28	0.78
Epi-retinal Membrane	269	82	0.49	0.37	0.61
Epi-retinal Membrane with foveal deformation	269	90	0.46	0.31	0.61
Intraretinal Fluid (Cysts)	269	73	0.48	0.38	0.58
Intraretinal Fluid within fovea	269	73	0.48	0.38	0.58
Intraretinal Fluid at foveal center	269	85	0.63	0.53	0.73
Retinal Pigment Epithelial (RPE) Elevation	270	90	**	**	**
Sub-RPE Fluid	269	88	0.75	0.67	0.83
Sub-RPE Fluid within fovea	269	93	0.79	0.70	0.88
Sub-RPE Fluid at foveal center	269	90	0.77	0.68	0.85
Subretinal Hyper Reflective Material	269	87	0.75	0.67	0.83
Subretinal Hyper Reflective Material within fovea	269	88	0.78	0.71	0.85
Subretinal Fluid	270	90	0.80	0.73	0.87
Subretinal Fluid within fovea	270	89	0.75	0.67	0.83
Subretinal Fluid at foveal center	270	93	0.81	0.72	0.89
Macular Fluid	270	84	0.55	0.43	0.67

OCT = optical coherence tomography, CL = confidence limit, Fovea = 1 mm diameter horizontal region centered at the middle of the fovea, VMA = any vitreous attachment and separation from the retina within a 3 mm diameter horizontal region centered at the middle of the fovea, foveal deformation = any retinal distortion within the fovea, RPE = retinal pigment epithelium, Macular fluid = presence of any one or more of the following: intraretinal fluid, subretinal fluid, or sub-RPE fluid.

* = A single scan had missing data during repeat grading of categorical measures.

** = kappa could not be computed for tables that were asymmetric or that were symmetric without matching rows and columns. The second grading team re-graded images at least 5 weeks after the images were initially graded by the first grading team. A morphological feature was not interpretable for grading if any one of the following conditions existed: inadequate positioning, poor saturation, absent scan, or portion of feature extends off radial line scan.

Table 2

Reading Center Team grading agreement for quantitative OCT evaluation

Variable	Statistic	Reading Center Team A	Reading Center Team B	Paired Difference	P-Value*
RPE Elevation Height - maximum vertical dimension on any scan (micrometers)					
N				226*	
Mean (SD)		165 (159)	177 (175)	-8 (42)	
Min.,Median,Max		22,121,1210	22,116,1210	-220,0,110	0.02
ICC (95% CI)				0.97 (0.96 to 0.97)	
RPE Elevation Width - maximum horizontal dimension on any scan (micrometers)					
N				210**	
Mean (SD)		771 (525)	808 (535)	-62 (313)	
Min.,Median,Max		110,616,2530	88,665,5,2607	-1320,-11,1760	0.001
ICC (95% CI)				0.81 (0.76 to 0.85)	
Mean Retinal Thickness at Foveal Center Point (micrometers)					
N				270	
Mean (SD)		174 (82)	177 (82)	-3 (30)	
Min.,Median,Max		18,161,557	11,161,559	-223,0,172	0.43
ICC (95% CI)				0.93 (0.92 to 0.95)	
Mean Subretinal Fluid Thickness at Foveal Center Point (micrometers)					
N				270	
Mean (SD)		10 (33)	9 (30)	0.6 (14)	
Min.,Median,Max		0,0,229	0,0,255	-37,0,167	0.33
ICC (95% CI)				0.90 (0.88 to 0.92)	
CNV-PED Complex Thickness at Foveal Center Point (micrometers)					
N				270	
Mean (SD)		144 (160)	145 (159)	-2 (30)	
Min.,Median,Max		22,68,1045	22,78,1027	-231,0,145	0.058
ICC (95% CI)				0.98 (0.98 to 0.99)	
Mean Total Thickness at Foveal Center Point (micrometers)					
N				270	
Mean (SD)		327 (181)	331 (182)	-4 (26)	
Min.,Median,Max		114,260,1136	112,264,1205	-84,0,99	0.026
ICC (95% CI)				0.99 (0.99 to 0.99)	

OCT = optical coherence tomography. RPE = retinal pigment epithelium. CNV-PED complex = Mean RPE + RPE Elevation + Subretinal Hyper Reflective Material thickness. ICC = intraclass correlation. SD = standard deviation. CI = confidence interval. Thickness refers to measurement of vertical dimension at the foveal center point unless otherwise specified. The second grading teams performed repeat

evaluation of images at least 5 weeks after the images were initially graded by the first grading team. A morphological feature was not interpretable for grading if any one of the following conditions existed: inadequate positioning, poor saturation, absent scan, or portion of feature extends off radial line scan.

* RPE Elevation not present on every scan for measurement.

** RPE Elevation width extended off screen preventing measurement in 16 scans.

Table 3

Reading Center Team qualitative agreement during four serial OCT evaluations.

Variable	Comparison between Grading Time	Total	Percent Agreement	Kappa Statistic	Kappa Statistic Lower 95% CL	Kappa Statistic Upper 95% CL
Vitreomacular attachment	1-2	23	100	1.00	1.00	1.00
	1-3	23	96	0.78	0.37	1.00
	1-4	23	100	1.00	1.00	1.00
Vitreomacular attachment with foveal deformation	1-2	23	100	1.00	1.00	1.00
	1-3	23	100	1.00	1.00	1.00
	1-4	23	100	1.00	1.00	1.00
Epiretinal Membrane	1-2	23	87	0.69	0.37	1.00
	1-3	23	74	*	*	*
	1-4	23	74	0.42	0.01	0.82
Epiretinal Membrane with foveal deformation	1-2	23	96	0.78	0.37	1.00
	1-3	23	83	*	*	*
	1-4	23	96	0.78	0.37	1.00
Intraretinal Fluid (Cysts)	1-2	23	65	0.35	-0.03	0.74
	1-3	23	57	*	*	*
	1-4	23	70	*	*	*
Intraretinal Fluid within fovea	1-2	23	65	0.38	0.04	0.73
	1-3	23	57	*	*	*
	1-4	23	65	*	*	*
Intraretinal Fluid at foveal center	1-2	23	91	0.79	0.51	1.00
	1-3	23	91	*	*	*
	1-4	23	83	0.57	0.19	0.95
Retinal Pigment Epithelial (RPE) Elevation	1-2	23	96	0.65	0.01	1.00
	1-3	23	100	1.00	1.00	1.00

Variable	Comparison between Grading Time	Total	Percent Agreement	Kappa Statistic	Kappa Statistic Lower 95% CL	Kappa Statistic Upper 95% CL
	1-4	23	96	*	*	*
Sub-RPE Fluid						
	1-2	23	78	*	*	*
	1-3	23	91	*	*	*
	1-4	23	91	0.83	0.60	1.00
Sub-RPE Fluid at foveal center	1-2	23	87	0.73	0.44	1.00
	1-3	23	87	0.73	0.47	0.98
	1-4	23	96	0.90	0.71	1.00
Sub-RPE Fluid within fovea	1-2	23	91	0.76	0.45	1.00
	1-3	23	87	0.64	0.32	0.97
	1-4	23	100	1.00	1.00	1.00
Subretinal Hyper Reflective Material	1-2	23	83	0.66	0.34	0.97
	1-3	22	82	0.65	0.33	0.97
	1-4	23	74	0.51	0.16	0.85
Subretinal Hyper Reflective Material within fovea	1-2	23	91	0.82	0.58	1.00
	1-3	23	87	0.74	0.46	1.00
	1-4	23	78	0.56	0.21	0.91
Subretinal Fluid						
	1-2	23	100	1.00	1.00	1.00
	1-3	23	91	0.81	0.54	1.00
	1-4	23	83	0.64	0.32	0.96
Subretinal Fluid within fovea	1-2	23	87	0.66	0.30	1.00
	1-3	23	83	0.57	0.20	0.95
	1-4	23	87	0.66	0.30	1.00
Subretinal Fluid at foveal center	1-2	23	100	1.00	1.00	1.00
	1-3	23	96	0.84	0.53	1.00
	1-4	23	100	1.00	1.00	1.00

Variable	Comparison between Grading Time	Total	Percent Agreement	Kappa Statistic	Kappa Statistic Lower 95% CL	Kappa Statistic Upper 95% CL
Macular Fluid	1-2	23	83	0.23	-0.30	0.77
	1-3	23	83	0.40	-0.07	0.87
	1-4	23	78	0.33	-0.11	0.77

OCT = optical coherence tomography. CL = confidence limit. Fovea = 1 mm diameter horizontal region centered at the middle of the fovea. VMA = any vitreous attachment and separation from the retina within a 3 mm diameter horizontal region centered at the middle of the fovea. Foveal deformation = any retinal distortion within the fovea. RPE = retinal pigment epithelium. Macular fluid = presence of any one or more of the following: intraretinal fluid, subretinal fluid, or sub-RPE fluid.

* = kappa could not be computed for tables that were asymmetric or that were symmetric without matching rows and columns. For this ongoing analysis, 23 randomly selected scans uploaded July 2009 were reevaluated by a repeat grading team at each of the following time points after baseline: time 1 = 3 months, time 2 = 8 months, and time 3 = 12 months. After repeat grading, values were compared to those from initial grading. A repeat grading team was comprised of a Senior Reader and two readers who had not initially graded a particular scan. A morphological feature was not interpretable for grading if any one of the following conditions existed: inadequate positioning, poor saturation, absent scan, or portion of feature extends off radial line scan.

Table 4
Reading Center Team quantitative agreement during four serial OCT evaluations

Variable	Times	Statistic	Value at Time 1	Value at Repeat Grading	Paired Difference	P-Value *	Intraclass Correlation (95% CI)
RPE Elevation Height - maximum vertical dimension on any scan							
	1-2	N			21 **		
		Mean (SD)	122 (84)	134 (89)	-7 (26)	0.215	0.95 (0.89, 0.98)
		Min.,Median,Max	22,93,407	55,99,440	-55,0,44		
	1-3	N			22 **		
		Mean (SD)	122 (84)	133 (91)	-11 (23)	0.021	0.96 (0.90, 0.98)
		Min.,Median,Max	22,93,407	33,104,418	-55,-11,44		
	1-4	N			22 **		
		Mean (SD)	122 (84)	135 (81)	-12 (26)	0.041	0.94, 0.87, 0.98
		Min.,Median,Max	22,93,407	55,110,396	-66,-16,44		
RPE Elevation Width - maximum horizontal dimension on any scan							
	1-2	N			20 **		
		Mean (SD)	730 (520)	898 (646)	-146 (488)	0.237	0.64 (0.29, 0.84)
		Min.,Median,Max	110,616,2068	110,880,2310	-1320,-27,605		
	1-3	N			20 **		
		Mean (SD)	730 (520)	749 (461)	-32 (331)	0.821	0.78 (0.53, 0.90)
		Min.,Median,Max	110,616,2068	88,632,1595	-880,-5,495		
	1-4	N			21 **		
		Mean (SD)	730 (520)	825 (495)	-108 (325)	0.151	0.78 (0.54, 0.90)
		Min.,Median,Max	110,616,2068	121,638,1760	-968,-66,528		
Mean Retinal Thickness at Foveal Center Point (micrometers)							
	1-2	N			23		
		Mean (SD)	162 (50)	162 (46)	-0.2 (13)	0.880	0.97 (0.92, 0.99)
		Min.,Median,Max	48,148,304	57,150,302	-26,0,26		
	1-3	N			23		
		Mean (SD)	162 (50)	163 (46)	-0.7 (12)	0.757	0.97 (0.93, 0.99)

Variable	Times	Statistic	Value at Time 1	Value at Repeat Grading	Paired Difference	P-Value*	Intraclass Correlation (95% CI)
		Min,Median,Max	48,148,304	55,156,299	-22,-2,20		
	1-4	N			23		
		Mean (SD)	162 (50)	165 (45)	-0.7 (12)	0.757	0.95 (0.90, 0.98)
		Min,Median,Max	48,148,304	71,152,304	-22,-2,20		
Mean Subretinal Fluid Thickness at Foveal Center Point (micrometers)							
	1-2	N			23		
		Mean (SD)	4 (13)	4 (13)	-0.1 (0.4)	1.000	1.00 (1.00, 1.00)
		Min,Median,Max	0,0,60	0,0,60	-2,0,0		
	1-3	N			23		
		Mean (SD)	0,0,60	4 (13)	-0.4 (3)	1.000	0.98 (0.94, 0.99)
		Min,Median,Max	4 (13)	0,0,57	-13,0,4		
	1-4	N			23		
		Mean (SD)	4 (13)	4 (13)	0 (0)		1.00
		Min,Median,Max	0,0,60	0,0,60	0,0,0		
CNV-PED Complex Thickness at Foveal Center Point (micrometers)							
	1-2	N			23		
		Mean (SD)	103 (102)	106 (102)	-2 (21)	0.730	0.98 (0.95, 0.99)
		Min,Median,Max	33,49,433	24,59,423	-48,0,28		
	1-3	N			23		
		Mean (SD)	103 (102)	112 (101)	-9 (22)	0.108	0.97 (0.94, 0.99)
		Min,Median,Max	33,49,433	31,60,434	-79,-9,16		
	1-4	N			23		
		Mean (SD)	103 (102)	104 (99)	-0.9 (22)	0.935	0.98 (0.95, 0.99)
		Min,Median,Max	33,49,433	31,53,412	-51,4,37		
Mean Total Thickness at Foveal Center Point (micrometers)							
	1-2	N			23		
		Mean (SD)	269 (113)	272 (115)	-3 (28)	0.713	0.97 (0.94, 0.99)
		Min,Median,Max	160, 234, 546	149, 231, 578	-73,0,35		
	1-3	N			23		

Variable	Times	Statistic	Value at Time 1	Value at Repeat Grading	Paired Difference	P-Value*	Intraclass Correlation (95% CI)
		Mean (SD)	270 (113)	279 (110)	-10 (25)	0.092	0.97 (0.94, 0.99)
		Min,Median,Max	160, 234, 546	161, 233,563	-83,-9,,26		
	1-4	N			23		
		Mean (SD)	269 (113)	273 (106)	-4 (26)	0.778	0.97 (0.94, 0.99)
		Min,Median,Max	160, 234, 546	172, 226, 546	-70,0,29		

OCT = optical coherence tomography. RPE = retinal pigment epithelium. CNV-PED complex = Mean RPE + RPE Elevation + Subretinal Hyper Reflective Material thickness. CI = confidence interval. For this ongoing analysis, 23 randomly selected scans uploaded July 2009 were reevaluated by a repeat grading team at each of the following time points after baseline: time 1 = 3 months, time 2 = 8 months, and time 3 = 12 months. After repeat grading, values were compared to those from initial grading. A repeat grading team was comprised of a Senior Reader and two readers who had not initially graded a particular scan. RPE = retinal pigment epithelium.

* P-value based on Wilcoxon signed rank test of median difference equal to zero.

** RPE elevation not present on every scan for measurement.