



OPEN ACCESS

# Assessment of estimated retinal atrophy progression in Stargardt macular dystrophy using spectral-domain optical coherence tomography

Rupert W Strauss,<sup>1,2,3</sup> Beatriz Muñoz,<sup>1</sup> Yulia Wolfson,<sup>1</sup> Raafay Sophie,<sup>1</sup> Emily Fletcher,<sup>1,4</sup> Millena G Bittencourt,<sup>1</sup> Hendrik P N Scholl<sup>1</sup>

► Additional material is published online only. To view please visit the journal online (<http://dx.doi.org/10.1136/bjophthalmol-2015-307035>).

<sup>1</sup>The Wilmer Eye Institute, Johns Hopkins University, Baltimore, USA

<sup>2</sup>Department of Ophthalmology, Medical University Graz, Graz, Austria

<sup>3</sup>Department of Ophthalmology, General Hospital Linz, Linz, Austria

<sup>4</sup>Department of Ophthalmology, Cheltenham General Hospital, Gloucestershire Hospitals NHS Foundation Trust, Cheltenham, UK

## Correspondence to

Professor Hendrik P N Scholl, Wilmer Eye Institute, Johns Hopkins University School of Medicine, 748 Maumenee Building, 1800 Orleans St, Baltimore, MD 21287-9277, USA; [hscholl1@jhmi.edu](mailto:hscholl1@jhmi.edu)

Received 16 April 2015

Revised 30 September 2015

Accepted 9 October 2015

Published Online First

14 November 2015



Open Access  
Scan to access more  
free content

## ABSTRACT

**Aims** To estimate disease progression based on analysis of macular volume measured by spectral-domain optical coherence tomography (SD-OCT) in patients affected by Stargardt macular dystrophy (STGD1) and to evaluate the influence of software errors on these measurements.

**Methods** 58 eyes of 29 STGD1 patients were included. Numbers and types of algorithm errors were recorded and manually corrected. In a subgroup of 36 eyes of 18 patients with at least two examinations over time, total macular volume (TMV) and volumes of all nine Early Treatment of Diabetic Retinopathy Study (ETDRS) subfields were obtained. Random effects models were used to estimate the rate of change per year for the population, and empirical Bayes slopes were used to estimate yearly decline in TMV for individual eyes.

**Results** 6958 single B-scans from 190 macular cube scans were analysed. 2360 (33.9%) showed algorithm errors. Mean observation period for follow-up data was 15 months (range 3–40). The median (IQR) change in TMV using the empirical Bayes estimates for the individual eyes was  $-0.103$  ( $-0.145$ ,  $-0.059$ )  $\text{mm}^3$  per year. The mean ( $\pm$ SD) TMV was  $6.321 \pm 1.000$   $\text{mm}^3$  at baseline, and rate of decline was  $-0.118$   $\text{mm}^3$  per year ( $p=0.003$ ). Yearly mean volume change was  $-0.004$   $\text{mm}^3$  in the central subfield (mean baseline= $0.128$   $\text{mm}^3$ ),  $-0.032$   $\text{mm}^3$  in the inner (mean baseline= $1.484$   $\text{mm}^3$ ) and  $-0.079$   $\text{mm}^3$  in the outer ETDRS subfields (mean baseline= $5.206$   $\text{mm}^3$ ).

**Conclusions** SD-OCT measurements allow monitoring the decline in retinal volume in STGD1; however, they require significant manual correction of software errors.

## INTRODUCTION

Stargardt macular dystrophy (STGD1) is a predominantly autosomal-recessive inherited retinal disorder due to pathogenic mutations in the *ABCA4* gene.<sup>1</sup> STGD1 is the most common cause of juvenile macular degeneration (estimated prevalence 10–12.5 per 100 000) and typically shows progressive impairment of visual acuity, which begins most frequently within the first or second decades of life.<sup>2</sup> The loss of acuity is accompanied by atrophic-appearing lesions within the macula and the presence of yellow-white lesions at the level of the retinal pigment epithelium (RPE), which are referred to as ‘fundus flecks’ caused by excess accumulation of lipofuscin in the RPE.<sup>3</sup> It is characterised by a degeneration of photoreceptors that occurs as a result of the formation of toxic bisretinoids in photoreceptor outer segments or

secondarily following loss of the RPE due to lipofuscin (bisretinoid) accumulation.<sup>3</sup>

So far, there is no therapy for this disease. However, different therapeutic approaches are in the early clinical phase including stem cell therapy,<sup>4</sup> gene therapy (ClinicalTrials.gov identifier NCT01736592, NCT01367444) and medical therapy that markedly slow the formation of toxic components (ClinicalTrials.gov identifier NCT02402660). Any new therapeutic option must be compared with the natural course of a disease, and for surrogate measures, the US Food and Drug Administration states that “validated surrogate markers are those for which evidence has been established that a drug-induced effect on the surrogate predicts the desired effect on the clinical outcome of interest”.<sup>5</sup>

Remarkable progress has been made in high-resolution imaging of the retina using new technologies such as spectral-domain optical coherence tomography (SD-OCT).<sup>6</sup> Using SD-OCT, visualisation of intraretinal architectural morphology has been achieved that previously had been possible only with histopathology<sup>7</sup> and which may serve as possible clinical outcome measures. Furthermore, SD-OCT imaging does not only qualitatively reveal the structural changes that occur in macular diseases, but also provides quantitative measurements of the means of retinal thickness and volume.<sup>7</sup> Indeed, inclusion criteria for clinical trials and retreatment protocols in trials of intravitreal agents for macular disease like diabetic macular oedema often incorporate a quantitative OCT parameter based on the nine Early Treatment of Diabetic Retinopathy Study (ETDRS) subfield grid.<sup>8,9</sup> Furthermore, the concept of measuring SD-OCT-derived retinal measures, such as peripapillary retinal nerve fibre layer thickness and total macular volume (TMV), have been proposed as outcome measures in remyelinating and neuroprotective trials in multiple sclerosis.<sup>10</sup>

The Spectralis OCT (Heidelberg Engineering, Heidelberg, Germany) offers the opportunity to follow changes longitudinally by using a system that features two different options: an eye tracker that can compensate for eye movements during scanning procedure and can improve reproducibility of the measurements made, and a retest function that ensures that follow-up measurements are performed at the same retinal locus as baseline examination.<sup>11</sup> However, image errors still occur in the majority of SD-OCT scans in different macular diseases despite various advances of the SD-OCT technology compared with the former time-domain



**To cite:** Strauss RW, Muñoz B, Wolfson Y, et al. *Br J Ophthalmol* 2016;**100**:956–962.

OCT, such as improved image quality, faster image acquisition and increased sensitivity of SD-OCT.<sup>9</sup> We hypothesised that with manual correction of such errors macular volume may be an efficient method to detect retinal atrophy. Therefore, the purpose of this study was to estimate disease progression based on analysis of macular volume measured by Spectralis SD-OCT in patients affected by STGD1 in individual clinics and to evaluate the susceptibility of the Spectralis OCT to errors.

## PATIENTS AND METHODS

Diagnosis of STGD1 was based on all of the following criteria:<sup>12–14</sup> (1) fundus flecks at the level of the RPE typical for STGD1, (2) RPE changes and/or bull's eye appearance and/or atrophy of the RPE and (3) a familial history compatible with recessive inheritance pattern. In 14 patients, the diagnosis was molecularly confirmed by genetic testing for pathogenic mutations in the *ABCA4* gene by CLIA-certified laboratories through the EyeGene project (ClinicalTrials.gov identifier: NCT00378742). Three patients showed one mutation and 11 patients showed two mutations in the *ABCA4* gene (further information is provided in the online supplementary material). Fundoscopic appearance was used for classification according to Fishman *et al.*<sup>15</sup> group 1 was characterised by a central, atrophic-appearing macular lesion with or without flecks; patients of group 2 showed numerous flecks, extending nasally to the optic nerve head and temporal to the vascular arcades; and group 3 was defined by extensive atrophic-appearing RPE changes and partially resorbed flecks.

Inclusion criteria for this study were the diagnosis of STGD1 based on the above-mentioned criteria, clear optical media and presence of at least one Spectralis SD-OCT cube scan per eye. Exclusion criteria were (1) presence of macular diseases other than STGD1, (2) presence of macular oedema, (3) absence of any atrophic alterations of the RPE, (4) opacity of optical media, (5) ungradable images of low quality due low signal strength or uncorrectable errors like motion artefacts or severely degraded images and (6) scans of eyes with atrophic lesions that exceed >80% of the ETDRS grid in the output of the software. Eyes also were excluded if after three attempts of acquisition such errors could not be overcome, for example, due to patient's inability to maintain stable fixation. Baseline characteristics of the patients included age at first examination, gender, race and age at onset of disease. All patients underwent a full clinical examination including testing of best-corrected visual acuity (BCVA) and SD-OCT as part of their routine clinical care. BCVA was measured using Snellen visual acuity charts and were converted into logarithm of the minimum angle of resolution for statistical purposes.

## Examinations

Pupils were dilated for SD-OCT using one drop 1% tropicamide and 2.5% neosynephrine. All measurements were taken by one experienced OCT technician (DE) certified for obtaining clinical trial images for reading centres using a Spectralis OCT instrument running HRA2/Spectralis Family Acquisition Module (Heidelberg Engineering). Patients were controlled for correct sitting position, position of the head, focus of the video image and centralisation of the scan area before acquisition was started. Scans were obtained in the high-speed mode with the automated real-time mode activated and set to 5 (creating a mean image of five repeated identical B-scans to improve signal-to-noise ratio). Centred on the fovea, one of three different cube scan protocols, containing a different density of B-scan lines (19, 25 or 49, respectively), was chosen by the discretion

of the photographer, covering an area of 20°×20° (for 25 or 49 B-scan lines, respectively) or 20°×15° (for 19 B-scan lines). For stabilisation of fixation, an internal or external fixation light (in case the patient was not able to see the internal one) was used. The first examination was marked as the reference, follow-up examinations as the follow-up, and the same setting and position of the scans were applied. During the last visit, one or both of the other two acquisition protocols were performed in a subgroup of patients in addition to the follow-up examination.

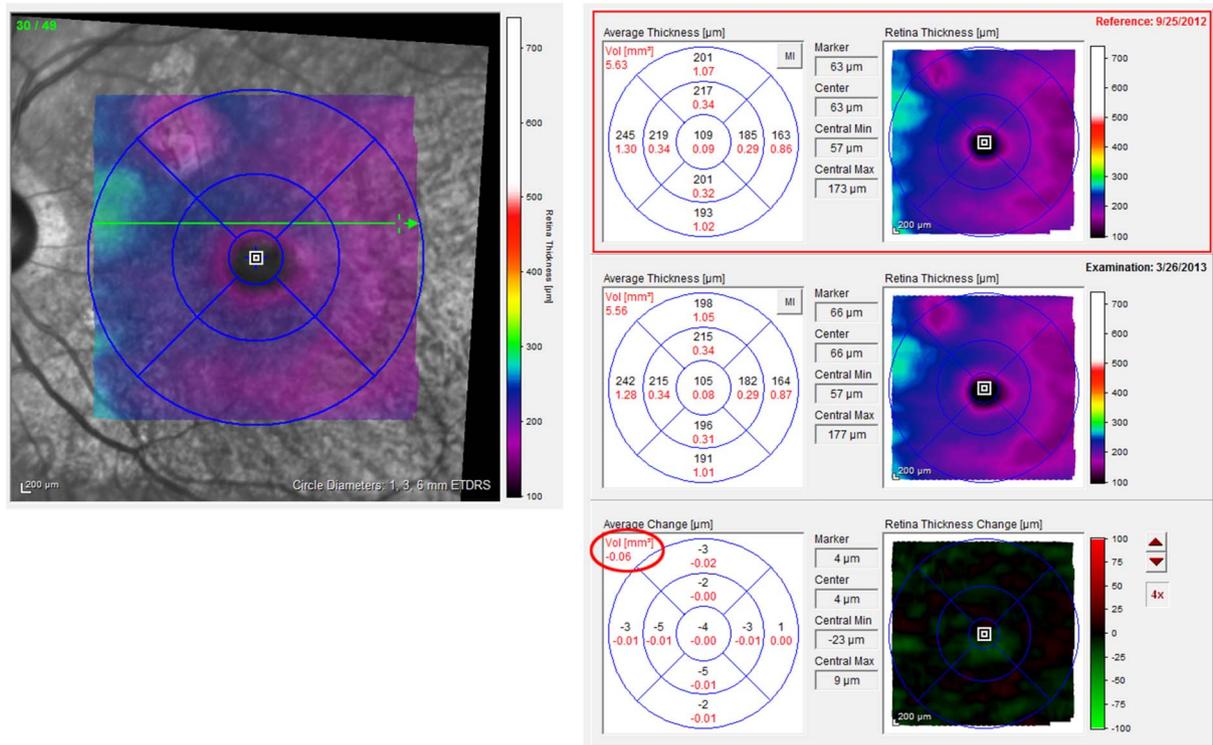
## Evaluation of segmentation errors

The standard output of the software version (HRA/Spectralis Viewing Module, V5.3.2.0; Heidelberg Engineering) offers the possibilities to display a false colour topographical as well as a thickness map showing measurements and providing values of retinal thickness and volume calculated for each of the nine macular areas corresponding to the ETDRS subfield grid<sup>16</sup>: this consists of a central circle (diameter 1 mm) and two surrounding rings: the inner is formed by two circles with diameters of 1 and 3 mm, respectively, and the outer ring of two circles with diameters of 3 and 6 mm, respectively. The inner and outer rings are divided into four equally large areas. The software also allows reviewing each B-scan individually and provides a difference map to compare changes over time (figure 1). The automatically set threshold algorithm sets the anterior line at the inner limiting membrane and the posterior line at the posterior surface of the third hyperreflective line of the retina corresponding to Bruch's membrane according to manufacturer's specifications.<sup>17</sup> The total number of B-scans per volume scan was registered and each single B-scan was reviewed by one examiner (RWS) to determine whether the automatically set threshold algorithm lines were set correctly. Errors of the algorithm (figure 2) were recorded, graded and corrected according to previously published criteria.<sup>9 17 18</sup> Details about the frequency of different kind of segmentation errors, the influence of the errors on measurement outcomes and the comparability of different B-scan densities are provided in the online supplementary material. For this purpose, the kind and the horizontal length of each error in each line was registered and an error score was calculated, specifying the accuracy of line segmentation and measurement: sum of error lengths/sum of total lines lengths of each macular cube. The depths of errors affecting the central five subfields were recorded. In a subgroup of 21 patients (42 eyes), the values of retinal volume before and after the correction of segmentation errors were compared. During the last visit, two separate scans were acquired with two different scans densities (25 and 49 B-scans per cube, respectively) in a subgroup of 10 patients (20 eyes). The segmentations errors were corrected and the comparability was assessed in order to determine whether a lower scan density can provide the same information about TMV as this would facilitate the application of using SD-OCT to track retinal atrophy in routine clinical practice.

## Evaluation of progression of retinal atrophy

After correction of all correctible segmentation errors, the values of volume measurements for each of the nine ETDRS subfields as well as TMV of all patients were transferred to a Microsoft Excel (Microsoft, Redmond, Washington, USA) spreadsheet for further analysis (figure 1). Date of examination and image acquisition were noted.

For 36 eyes of 18 patients with images available for at least two separate visits, retinal volumes of all 9 ETDRS subfields as well as TMV were registered for each visit. In case of more than



**Figure 1** Example of the standard output of the Spectralis optical coherence tomography (OCT) measurements map using the Early Treatment of Diabetic Retinopathy Study (ETDRS) grid for different follow-up examinations after manual correction of segmentation errors. Left side: confocal scanning laser ophthalmoscope image with scanned area (macular cube in false colour code) is provided. Right side: (from top to bottom) measurements of retinal volume (red numbers) and total macular volume (TMV, red circle) and all nine ETDRS subfields are given for the first examination (reference) and the follow-up examination (middle) including retinal and volume. The differences for these measurements are given in the average change map (bottom). In this example, a decrease in TMV of  $-0.06 \text{ mm}^3$  was observed.

two consecutive visits were available for the patient, each following visit was registered to the former one. Differences between visits were calculated and rate of change per time (loss of retinal volume per year) estimated. These changes were also calculated for the inner circle (mean retinal volume of the inner four subfields) and outer circle (mean retinal volume of the outer four subfields). Furthermore, progression rates of loss of TMV were also estimated for subgroups of patients with confirmed mutations in the *ABCA4* gene and for subgroup of patients after classified according to Fishman *et al*<sup>15</sup> (see online supplementary material). To estimate the coefficient of repeatability (CR), the scan procedure and the measurements of macular volume were repeated during the last visit (using the eye-tracker option and ‘follow-up’ mode) in 20 eyes of 10 patients with follow-up data after manual correction of algorithm errors in all B-scans. The CR is the value below which the absolute differences between two measurements would lie with 0.95 probability. It was calculated as 2.77 times the within eyes SD.<sup>19</sup>

**Statistical analysis**

Linear mixed models, with time as the independent variable, were used to estimate the yearly change for each outcome. The models include random effects for the intercept, the slope for time and the eye. This modelling approach incorporates not only the statistical dependencies of within-individuals’ repeated assessments over time but also the between eyes expected correlation. Specifically, the outcome at time *t* on the *j*th eye of the *I*th individual was modelled as  $Y_{ijt} = \alpha + \beta \times t + a_i + b_i \times t + c_j + e_{ijt}$  where (*a<sub>i</sub>*, *b<sub>i</sub>*) followed a bivariate normal with zero means, *c<sub>j</sub>* followed a normal distribution with zero mean and variance

corresponding to the between eyes variability and the *e<sub>ijt</sub>* are the within-individuals’ departures of the observed data from the individuals’ lines. For individual eyes, empirical Bayes slopes were calculated to estimate yearly decline in TMV and volumes in each of the ETDRS subfields. The results of the random effects model were used in order to achieve a weighted average of the empirical slope based on individual data and the average of the population as a whole. The Wilcoxon test for related samples was used to analyse differences before and after manual correction and the Spearman’s *r* for independent samples to analyse correlations between TMV and error score. Statistical significance was defined as  $p < 0.05$ . Statistical analyses were conducted using SPSS software (V21.0; SPSS, Chicago, Illinois, USA), SAS V9.3. (Cary, North Carolina, USA) and R project for statistical computing (V2.15.1, Institute for Statistics and Mathematics, Vienna).

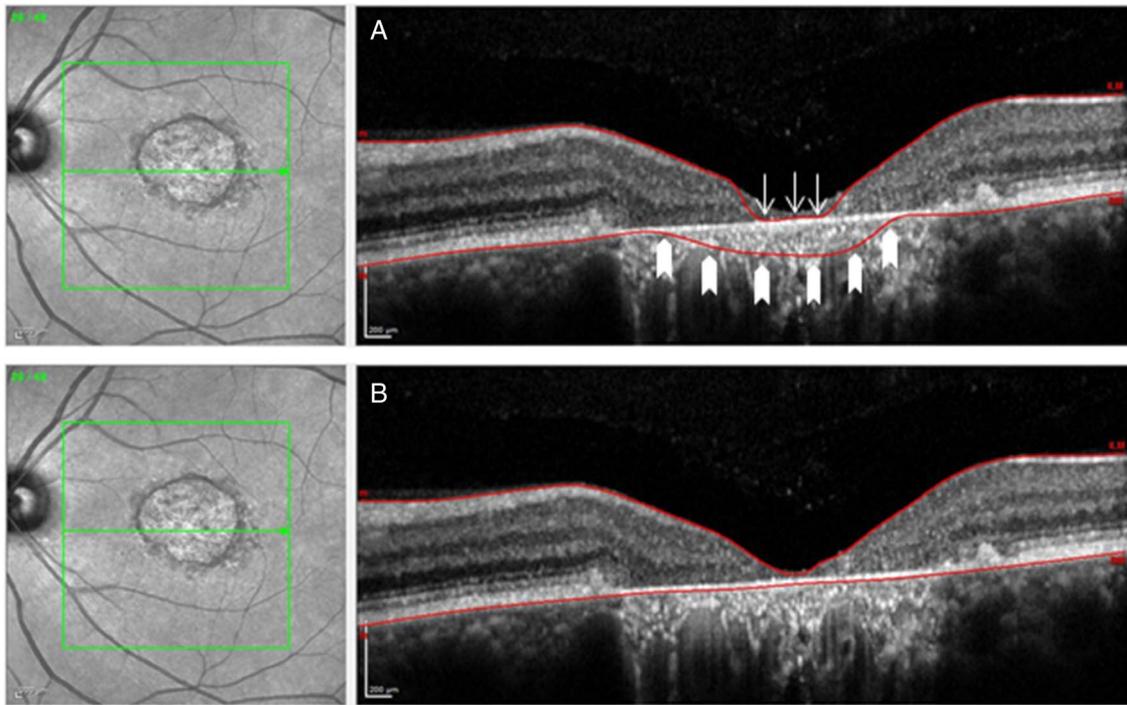
**RESULTS**

**Demographics**

Images of 58 eyes of 29 patients affected with STGD1 were included in the analysis. Demographic data of all patients and patients with available follow-up visits are summarised in table 1.

**Evaluation of algorithm errors**

Detailed information about the evaluation of algorithm errors is provided in the online supplementary material. Algorithm errors were detected and corrected in 2360 of 6958 reviewed B-scans (33.9%). There was a negative correlation between TMV and error score for the two scan densities of 19 ( $n=30$  eyes,  $r_s = -0.413$ ,  $p=0.023$ ) and 49 B-scans per cube ( $n=44$ ,



**Figure 2** Confocal scanning laser ophthalmoscope image on the left side with the corresponding cross-sectional single B-scan on the right side, as provided by the software. Examples exhibit algorithm errors prior to (A) and after (B) manual correction of these segmentation errors. The example in (A) shows both a misidentification of the inner retina (arrows) and the outer (block arrows) retinal limits, respectively.

$r_s = -0.513$ ,  $p = 0.000$ ), respectively, but not for the scan density (25 B-scans per cube scan) with the lowest number of eyes ( $n = 24$  eyes,  $r_s = -0.188$ ,  $p = 0.378$ ).

For all three different scan densities, there was a statistically significance difference ( $p < 10^{-3}$ , respectively) in TMV before and after manual correction of segmentation errors in the investigated subset of 42 eyes (21 patients).

### Yearly change of retinal volume

Follow-up data of at least two visits (baseline and follow-up visit) were available for 36 eyes of 18 subjects (range 1–4 follow-up visits). Results for the TMV as indicated by the

output number of the software, as well as volumes in all nine single ETDRS-subfields at time of first visit of these 18 patients, are presented in [table 2](#). This table also provides the estimated yearly change of retinal volume based on different follow-up periods (mean 15 months, range 3–40 months). The mean ( $\pm$ SD) TMV was  $6.321 \pm 1.000 \text{ mm}^3$  at baseline and the rate of decline was  $-0.118 \text{ mm}^3$  per year ( $p = 0.003$ ; see [figure 3](#)). Mean volume change was  $-0.004 \text{ mm}^3$  per year in the central subfield (mean baseline =  $0.128 \text{ mm}^3$ ),  $-0.032 \text{ mm}^3$  per year in the inner ETDRS circle (inner four subfields; mean baseline =  $1.484 \text{ mm}^3$ ) and  $-0.079 \text{ mm}^3$  per year in the outer

**Table 1** Demographics of all patients included in the study and of the subgroup of patients with at least two visits (follow-up)

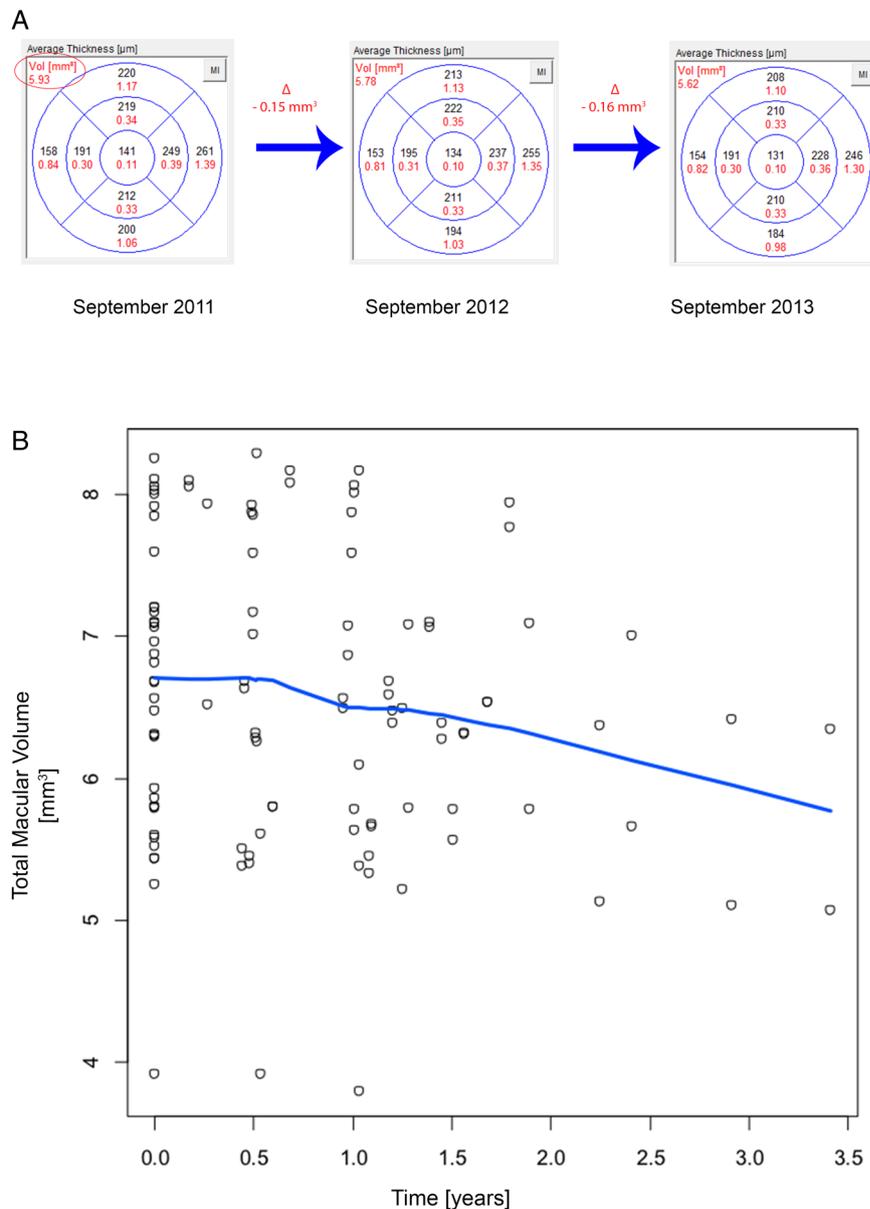
	All 29 patients included in the study	Patients with follow-up visits (n=18)
Median age (range (years)) of onset of the disease	30.0 (10.0–68.0)	36.0 (11.0–68.0)
Median age (range (years)) at first visit	37.1 (16.5–71.0)	43.5 (23.3–71.0)
Median (range) of best corrected visual acuity (logarithm of the minimum angle of resolution) at first visit	0.35 (–1.0–1.3)	0.35 (–1.0–1.3)
Gender		
Male	16 (55.2%)	11 (61.1%)
Female	13 (44.8%)	7 (38.9%)
Race		
Caucasian	19 (65.5%)	13 (72.2%)
Afro-American	6 (20.7%)	4 (22.2%)
Asian	2 (6.9%)	0 (0%)
Indian	2 (6.9%)	1 (5.6%)

**Table 2** Retinal volume(s) in 36 eyes of 18 patients at baseline

Location	Baseline Mean (SD) ( $\text{mm}^3$ )	Estimated yearly change (95% CI) ( $\text{mm}^3$ )
Total macular volume	6.621 (1.000)	$-0.118$ ( $-0.174$ to $-0.049$ )
Central subfield	0.128 (0.048)	$-0.004$ ( $-0.007$ to $-0.002$ )
Inner circle (inner four subfields)		
Total	1.484 (0.237)	$-0.032$ ( $-0.045$ to $-0.020$ )
Superior	0.391 (0.057)	$-0.006$ ( $-0.010$ to $-0.001$ )
Nasal	0.373 (0.064)	$-0.009$ ( $-0.012$ to $-0.006$ )
Inferior	0.351 (0.064)	$-0.011$ ( $-0.015$ to $-0.007$ )
Temporal	0.370 (0.064)	$-0.009$ ( $-0.012$ to $-0.005$ )
Outer circle (outer four subfields)		
Total	5.206 (0.698)	$-0.079$ ( $-0.136$ to $-0.022$ )
Superior	1.355 (0.193)	$-0.017$ ( $-0.035$ to $-0.002$ )
Nasal	1.374 (0.175)	$-0.024$ ( $-0.036$ to $-0.013$ )
Inferior	1.216 (0.192)	$-0.018$ ( $-0.036$ to $0.000$ )
Temporal	1.258 (0.174)	$-0.021$ ( $-0.035$ to $-0.006$ )

Based on different follow-up periods (3–40 months), the estimated yearly changes were calculated for total macular volume, volumes in all single Early Treatment of Diabetic Retinopathy Study subfields and the inner and outer circle.

**Figure 3** The change of total macular volume (TMV) is shown over time. (A) Examples of the measurements (retinal thickness and volume (red)) in all Early Treatment of Diabetic Retinopathy Study (ETDRS) of the right eye of one patient are given at three different visits in a yearly interval. The decrease (delta) of TMV (red circle) between the visits is illustrated. (B) Pooled data of all patients with follow-up visits are shown. The estimated change in of TMV is graphically shown (blue line). The mean ( $\pm$ SD) TMV was  $6.321 \pm 1.000 \text{ mm}^3$  at baseline and the rate of decline was  $-0.118 \text{ mm}^3$  per year ( $p=0.003$ ).



ETDRS subfields (mean baseline= $5.206 \text{ mm}^3$ ). The median (IQR) change in TMV using the empirical Bayes estimates for the individual eyes was  $-0.103$  ( $-0.145, -0.059$ )  $\text{mm}^3$  per year, and in 14% of the eyes the estimated decline was  $>0.2 \text{ mm}^3$  per year. Yearly changes for subgroups of patients with genetically confirmed mutations in the *ABCA4* gene and for the subgroup of patients following the clinical phenotype classification according to Fishman are provided in the online supplementary material.

**DISCUSSION**

Neurodegeneration of the retina can be monitored by commercially available retinal imaging devices accessible in clinical setting.<sup>20</sup> The Spectralis OCT offers several features to monitor atrophic processes in macular disease such as STGD1: an eye tracker can compensate for eye movements during the scanning procedure, and the retest function places follow-up examinations at the same retinal locus as in the baseline examination. The Spectralis software sets the posterior boundary line not only in the area of the third hyperreflective band of the outer retina like other SD-OCT machines, but even more posteriorly

at the level of Bruch’s membrane including the RPE and therefore resulting in higher values of retinal thickness and volume even in normal eyes.<sup>17 21</sup> The inclusion of the RPE in retinal thickness and volume measurements may be of advantage to monitor atrophic processes in STGD1: the loss of tissue beginning at the level of the RPE or photoreceptors should be detectable by SD-OCT measurements using TMV as an outcome measure; on the other hand, however, Sparrow *et al*<sup>22</sup> could show in a recent publication that the bright short-wave length autofluorescence signal of flecks likely originates from augmented lipofuscin formation in degenerating photoreceptor cells impaired by the failure of RPE, and RPE cells associated with flecks in STGD1 are considerably changed or lost. To which extent the increase of volume is due to newly developed lipofuscin or other material, and the decrease is due to cell death or the noise of the software remains to be determined and will require segmentation analysis of retinal layers including the RPE.

In this study, we observed a statistically significant decline in TMV over time and could also compute a yearly decrease in retinal volume for all nine ETDRS subfields. Studies on

repeatability of the Spectralis OCT measurements of TMV are scant; to date, there are no reports specific to subjects with STGD1. Using data from 50 eyes and 4 repeated measures in eyes with diabetic macular oedema, Comyn *et al*<sup>8</sup> report a CR of 0.21 mm<sup>3</sup> for the total TMV. However, these authors did not correct the errors and only excluded scans with significant boundary detection errors. In our study, using two repeated measures in 20 eyes, after performing manual correction of algorithm errors in all B-scans, the CR estimate was 0.04 mm<sup>3</sup>, much lower than the previously reported. Based on the empirical Bayes estimate of the slope, 14% of the eyes in this study have a rate of decline of >0.20 mm<sup>3</sup> per year; to be conservative, we would label this group as ‘having true progression in a 1-year period’. Analogously, we estimate that 50% of the eyes will show signs of true decline in macular volume in a 2-year period. Using the above parameters, a hypothetical interventional 1-year trial to arrest the decline in TMV will require 328 eyes in each arm to show a 50% reduction in true progression with 80% power a two-sided  $\alpha=0.05$ . The corresponding numbers with a similar design for a 2-year trial are 64 eyes in each arm. Based on these calculations, SD-OCT measurement should be considered as a monitoring tool in upcoming clinical trials. However, larger patient numbers would be necessary to estimate definite progression rates.

The use of retinal volume measurements as a follow-up parameter requires some further considerations: our data show that it is important to correct algorithm errors, like segmentation errors, for the use of SD-OCT; such corrections can be impossible in case of registration errors like out-of-register artefacts or degraded scan images. We observed these errors also in this study, but in all cases the errors occurred in the outermost B-scans. A careful review was undertaken to assure that these scans did not affect thickness and volume measurements, for example, were not covered by the ETDRS-grid even after replacement in this analysis.

The correction of most of the errors like misidentification of retinal boundaries or incomplete segmentation is possible; however, the procedure can be very time-consuming. The Heidelberg Eye Explorer software provides the possibility of little dots to be used to correct, align or replace a retinal boundary. The number of such dots can vary depending on the curvature and presence of angles in segmentation line, and can count up to  $\geq 35$  points in just one line that need to be corrected. Other available automated algorithms may facilitate this procedure.<sup>23</sup> Therefore, the intracorrelation coefficient between two algorithms, both covering an area of 20°×20° but with different density of scan lines (25 vs 49) was also calculated (see online supplementary material) and shows a good agreement, indicating that the comparison of TMV in patients imaged with different scan densities may be used for retrospective analyses. Although a higher density of B-scans may enhance the precision of the measurement and the Spectralis OCT even offers higher rates up to 521 scans, a lower scan density can be considered for routine clinical application. The correction of these algorithm errors is crucial to achieve reliable follow-up data as indicated by the significant difference in volume measurements prior to and after the manual correction. We observed that these errors occur more frequently in atrophic regions (figure 2); this observation is reflected by the negative correlation between error score and TMV as an indicator of atrophic processes. Further studies are needed for the correlation between SD-OCT imaging and other imaging modalities like fundus autofluorescence (FAF) to evaluate which imaging modality may provide higher precision and therefore more power to detect atrophy progression.

Our study has several limitations: an observational period with a mean time period of 15 months may be too short to establish a definite progression rate for each individual patient. Second, we computed the yearly change based on pooled data from different patients that may show different progression rates based on their electrophysiological findings or genotype.<sup>24</sup> There are also reports about variable rates and speed of atrophy enlargement in patients with STGD1 as determined by the growth rate of atrophic lesions in FAF.<sup>12 25 26</sup> Third, although the concept to measure total retinal thickness and volume should reflect all changes in the retina and the RPE, the (sub) segmentation of the retinal layers and analysis of individual layers such as the outer nuclear layer or the outer segment length may provide a even more accurate progression rate than total retinal volume.<sup>23</sup>

Further improvement of the software tool is necessary to facilitate the measurements for the individual user. OCT-derived measurements of retinal thickness and volume may serve as outcome measures for clinical trials of clinical intervention that targets the progression of loss of the neuroretina and RPE in STGD1. Larger studies with a higher number of patients and a longer follow-up period are necessary to determine definite progression that may be dependent on genetic mutations or environmental factors like smoking or supplementation of vitamin A. Also, segmentation of individual retinal layers may provide higher precision and more power to detect small changes. Furthermore, the fundamental problem of functional consequences of structural abnormalities, such as loss of volume detected by high-resolution imaging, remain unclear. Therapeutic intervention in order to modify the retinal lesions detected with high-resolution imaging may only be justified if these abnormalities indeed cause visual loss. These questions are currently being addressed in the multicentre studies: “The natural history of the progression of atrophy secondary to Stargardt disease: a prospective and a retrospective longitudinal observational study” (ClinicalTrials.gov identifier NCT01977846).

**Acknowledgements** The authors thank David Emmert for expertise in acquisition of SD-OCT images and Kaoru Fujinami for advice in evaluation of genetic results.

**Contributors** RWS: design of the work, acquisition, analysis and interpretation of the data; drafting the manuscript. BM: analysis and interpretation of the data and revising the manuscript. YW: acquisition of the data and drafting the manuscript. RS: drafting the manuscript and revising the manuscript. EF: design of work, acquisition and analysis of the data and revising the manuscript. MGB: interpretation of the data and revising the manuscript. HPNS: design of the work, acquisition of the data, analysis and interpretation of the data, revising the manuscript. All authors approved the final version of the manuscript to be published and agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

**Funding** Erwin-Schroedinger Stipend by the Austrian Science Fund (Project # J3383-B23); Clinical/Research Fellowship Award in Inherited Orphan Retinal Degenerations of the Foundation Fighting Blindness (FFB); Wynn-Gund Translational Research Acceleration Program Enhanced Research and Clinical Training Award, National Neurovision Research Institute (NNRI)—Foundation Fighting Blindness (FFB); NNCD-CL-0310.0049-JHU-WG); Macular Degeneration Research Award, American Health Assistance Foundation (AHAF; M2010042); Unrestricted grant to the Wilmer Eye Institute from Research to Prevent Blindness; Baylor-Johns Hopkins Center for Mendelian Genetics (National Human Genome Research Institute, NHGRI/NIH; Identification number: 1U54HG006542-01). Supported in part by the Foundation Fighting Blindness Clinical Research Institute (FFB CRI) and a grant to FFB CRI by the U.S. Department of Defense USAMRMC TATRC, Fort Meade, Maryland (grant numbers W81-XWH-07-1-0720 and W81XWH-09-2-0189). H.P.N.S. is the Dr Frieda Derdeyn Bambas Professor of Ophthalmology. Dr Hendrik Scholl is consultant of the following entities (not including the National Institutes of Health): Fovea Pharmaceuticals; Trevena Inc.; Guidepoint Global, LLC; Gerson Lehrman Group. Dr Hendrik Scholl is member of the Data Monitoring Committee of the following entities (not including the National Institutes of Health): StemCells, Inc.; Genzyme Corp. Dr Hendrik Scholl serves as a member of the Ophthalmic Devices Panel of the

Medical Devices Advisory Committee, Food and Drug Administration (FDA), Silver Spring, Maryland, USA. Dr Hendrik Scholl is principal investigator of grants at The Johns Hopkins University sponsored by the following entity (not including the National Institutes of Health): QLT, Inc. Grants to investigators at The Johns Hopkins University are negotiated and administered by the institution (such as the School of Medicine) which receives the grants, typically through the Office of Research and Administration. Individual investigators who participate in the sponsored project(s) are not directly compensated by the sponsor but may receive salary or other support from the institution to support their effort on the project(s)

**Competing interests** None declared.

**Patient consent** Written consent was obtained from patients affected with STGD1 who were investigated at the retina division of the Wilmer Eye Institute, Johns Hopkins University, between 1 July 2009 and 1 October 2013.

**Ethics approval** Johns Hopkins School of Medicine Institutional Review Board. The study was conducted adhering to the tenets of the Declaration of Helsinki and in compliance with Health Insurance Portability and Accountability (HIPAA).

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Open Access** This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>

## REFERENCES

- Allikmets R, Singh N, Sun H, *et al*. A photoreceptor cell-specific ATP-binding transporter gene (ABCR) is mutated in recessive Stargardt macular dystrophy. *Nat Genet* 1997;15:236–46.
- Fishman GA, Farber M, Patel BS, *et al*. Visual acuity loss in patients with Stargardt's macular dystrophy. *Ophthalmology* 1987;94:809–14.
- Cideciyan AV, Aleman TS, Swider M, *et al*. Mutations in ABCA4 result in accumulation of lipofuscin before slowing of the retinoid cycle: a reappraisal of the human disease sequence. *Hum Mol Genet* 2004;13:525–34.
- Schwartz SD, Regillo CD, Lam BL, *et al*. Human embryonic stem cell-derived retinal pigment epithelium in patients with age-related macular degeneration and Stargardt's macular dystrophy: follow-up of two open-label phase 1/2 studies. *Lancet* 2015;385:509–16.
- Katz R. Biomarkers and surrogate markers: an FDA perspective. *NeuroRx* 2004;1:189–95.
- Drexler W, Fujimoto JG. State-of-the-art retinal optical coherence tomography. *Prog Retin Eye Res* 2008;27:45–88.
- Staurenghi G, Sadda S, Chakravarthy U, *et al*. Proposed lexicon for anatomic landmarks in normal posterior segment spectral-domain optical coherence tomography: the IN\*OCT consensus. *Ophthalmology* 2014;121:1572–8.
- Comyn O, Heng LZ, Ikeji F, *et al*. Repeatability of Spectralis OCT measurements of macular thickness and volume in diabetic macular edema. *Invest Ophthalmol Vis Sci* 2012;53:7754–9.
- Han IC, Jaffe GJ. Evaluation of artifacts associated with macular spectral-domain optical coherence tomography. *Ophthalmology* 2010;117:1177–89.e4.
- Warner CV, Syc SB, Stankiewicz AM, *et al*. The impact of utilizing different optical coherence tomography devices for clinical purposes and in multiple sclerosis trials. *PLoS ONE* 2011;6:e22947.
- Langenegger SJ, Funk J, Toteberg-Harms M. Reproducibility of retinal nerve fiber layer thickness measurements using the eye tracker and the retest function of Spectralis SD-OCT in glaucomatous and healthy control eyes. *Invest Ophthalmol Vis Sci* 2011;52:3338–44.
- McBain VA, Townend J, Lois N. Progression of retinal pigment epithelial atrophy in stargardt disease. *Am J Ophthalmol* 2012;154:146–54.
- Lois N, Holder GE, Bunce C, *et al*. Phenotypic subtypes of Stargardt macular dystrophy-fundus flavimaculatus. *Arch Ophthalmol* 2001;119:359–69.
- Fishman GA. Fundus flavimaculatus. A clinical classification. *Arch Ophthalmol* 1976;94:2061–7.
- Fishman GA, Stone EM, Grover S, *et al*. Variation of clinical expression in patients with Stargardt dystrophy and sequence variations in the ABCR gene. *Arch Ophthalmol* 1999;117:504–10.
- No authors listed]. Photocoagulation for diabetic macular edema. Early Treatment Diabetic Retinopathy Study report number 1. Early Treatment Diabetic Retinopathy Study research group. *Arch Ophthalmol* 1985;103:1796–806.
- Krebs I, Smreetschnig E, Moussa S, *et al*. Quality and reproducibility of retinal thickness measurements in two spectral-domain optical coherence tomography machines. *Invest Ophthalmol Vis Sci* 2011;52:6925–33.
- Sadda SR, Wu Z, Walsh AC, *et al*. Errors in retinal thickness measurements obtained by optical coherence tomography. *Ophthalmology* 2006;113:285–93.
- Vaz S, Falkmer T, Passmore AE, *et al*. The case for using the repeatability coefficient when calculating test-retest reliability. *PLoS ONE* 2013;8:e73990.
- Oberwahrenbrock T, Schippling S, Ringelstein M, *et al*. Retinal damage in multiple sclerosis disease subtypes measured by high-resolution optical coherence tomography. *Mult Scler Int* 2012;2012:530305.
- Wolf-Schnurrbusch UE, Ceklic L, Brinkmann CK, *et al*. Macular thickness measurements in healthy eyes using six different optical coherence tomography instruments. *Invest Ophthalmol Vis Sci* 2009;50:3432–7.
- Sparrow JR, Marsiglia M, Allikmets R, *et al*. Flecks in recessive stargardt disease: short-wavelength autofluorescence, Near-infrared autofluorescence, and optical coherence tomography. *Invest Ophthalmol Vis Sci* 2015;56:5029–39.
- Lazow MA, Hood DC, Ramachandran R, *et al*. Transition zones between healthy and diseased retina in choroideremia (CHM) and Stargardt disease (STGD) as compared to retinitis pigmentosa (RP). *Invest Ophthalmol Vis Sci* 2011;52:9581–90.
- Fujinami K, Lois N, Davidson AE, *et al*. A longitudinal study of stargardt disease: clinical and electrophysiologic assessment, progression, and genotype correlations. *Am J Ophthalmol* 2013;155:1075–88.e13.
- Chen B, Tosha C, Gorin MB, *et al*. Analysis of autofluorescent retinal images and measurement of atrophic lesion growth in Stargardt disease. *Exp Eye Res* 2010;91:143–52.
- Cukras CA, Wong WT, Caruso R, *et al*. Centrifugal expansion of fundus autofluorescence patterns in Stargardt disease over time. *Arch Ophthalmol* 2012;130:171–9.

# **Online only supplemental material: Evaluation and influence of algorithm errors of SD-OCT and yearly change in retinal volume for different subgroups of patients affected with Stargardt Macular Dystrophy (STGD1)**

Spectral-domain optical coherence tomography (SD-OCT) is widely used to monitor macular diseases. However, algorithm errors may significantly influence the measurements of retinal volume. One aspect of this study was therefore the evaluation and influence of algorithm errors on measurements in patients affected with STGD1.

## **Evaluation of algorithm errors**

In total, 190 macular cube scan consisting out of 6958 single B-scans were included in this analysis. Quality scores for scans are assigned by the Spectralis and expressed as a signal-to-noise ratio (SNR) in decibels (dB). Scans above 20 dB are considered high quality. [1] The average SNR value for the volume scans reported in this manuscript were  $27.4 \pm 5.3$  dB (mean  $\pm$  standard deviation).

Each single B-scan was reviewed by one examiner (RWS). The correct centration of the grid on the anatomical fovea was checked and adjusted if necessary. As the grid is centered by the photographer according to institution's standard operating procedures, a possible initial displacement of the grid by the software could not be evaluated.

The major types of errors that had been previously described in time-domain OCT and spectral domain OCT were evaluated [2] - these are illustrated in figure 2, supplemental figure 1 - and included: (1) misidentification of the inner retina (ILM), (2) misidentification of the outer retina, (3) degraded scan image, (4) out-of-register artifacts in which the inner retina was truncated because of a shifted scan, (5) inverted artifact in which the truncated superior and lateral edges of the retinal image were folded back onto the scan image in association with out-of-register artifacts, (6) complete segmentation error, (7) incomplete segmentation error and (8) cut edge

artifacts.

The horizontal length of each error in each line was registered and an error score was calculated specifying the accuracy of line segmentation and measurement (supplemental figure 2): sum of error lengths/sum of total lines lengths of each macular cube. For the last visit of each patient, the error score was correlated with the total macular volume (TMV).

Furthermore the depth was recorded in errors affecting the central 1-mm-diameter area and four inner subfields (supplemental figure 2). For this purpose, the scan lines forming the central and four inner ETDRS subfields were identified and these scans graded as stage 0 - no error; stage 1, as being an error of one third or less of the retinal thickness (mild error); stage 2 - error more than one third, but less than two thirds of the retinal thickness (moderate error); and stage 3 - more than two thirds of retinal thickness (severe error). Previous studies concentrated on the subevaluation of errors in the central subfield because this subfield reflects the visual function in term of visual acuity, and also can serve as an indicator for retreatment with intravitreal drug for macular edema. [3] In STGD1 however, macular edema is a rare complication and the aim of this study was to evaluate the atrophic process which typically occurs in a centrifugal pattern. [4] Therefore the parafoveal regions were added for analysis of center involving errors, too. Segmentation errors were corrected manually by the same examiner using the standard output software of the Spectralis™ until the misplaced segmentation lines were set at the correct position at the screen; missing segmentation lines were created, if necessary.

In total, 190 macular volume cube scans of 29 patients comprising in summary 6958 single B-scans were analyzed: 56 cube scans of 19 single B-scans (29.5%, 56/190), 28 cube scans of 25 single B-scans (14.8%, 28/190) and 106 cube scans of 49 single B-scans (55.7%, 106/190). 168 of these 190 macular cube scans (88.4%, 168/190) showed at least one type of segmentation error in at least one B-scan (mean 12.4 errors per macular cube scan, range 0 – 62). Supplemental table 1 summarizes the eight different types of segmentation errors in all 6958 reviewed single B-scans. Combined errors (e.g. at least two or more segmentation errors) were present in 282 (4.2%, 282/6958) single B-scans.

90 of the 190 macular cubes scans (47.4%, 90/190) revealed at least one center involving segmentation error; supplemental table 2 shows the frequency of center involving errors. Median error score (error length) of all B-scans was 0.03 (mean 0.12, range 0 – 0.78). There was negative correlation between error score and TMV for all three scan densities, when the last visit scans of all patients were analyzed. This was not statistically significant for the scan

density (25 B-scans per cube scan) with the lowest number of eyes ( $n = 24$  eyes,  $r_s = -0.188$ ,  $p = 0.378$ ), but showed statistical significance both for the scan densities of 19 ( $n = 30$  eyes,  $r_s = -0.413$ ,  $p = 0.023$ ) and 49 B-scans per cube ( $n = 44$ ,  $r_s = -0.5.13$ ,  $p = 0.000$ ). As some patients had imaging with several scan densities during their last visits, the sum exceeds the number of 58 eyes. Results are illustrated in supplemental figures 3.

### **Influence of correction of segmentation errors on values of total retinal volume**

In a subset of 21 patients (42 eyes), values of retinal volume in all nine subfields and TMV were transferred to a Microsoft Excel® (Microsoft Corp., Redmond, WA, U.S.A.) spreadsheet prior to correction of the segmentation errors during their last visit. Macular cube scans with a scan density of 19 and 25 B-scans were available for 18 eyes (9 patients) respectively and with a scan density of 49 macular cube scans for 32 eyes (16 patients). After correction of all segmentation errors, the values of retinal volume measurements for each of the nine ETDRS subfields as well as total TMV of all patients were transferred to a Microsoft Excel® (Microsoft Corp., Redmond, WA, U.S.A.) spreadsheet for further analysis.

For all three different scan densities, there was a statistical significance differences in the retinal volume before and after manual correction of segmentation errors in the investigated subset of 42 eyes (21 patients) as shown for TMV in supplemental figures 4.

### **Influence of algorithm errors**

The influence of errors and algorithm failures is of high importance. Some of these errors (i.e. misidentification of the inner or outer retinal border) can be manually corrected; others (i.e. out-of-register artifacts) are not correctible. Although Wolf-Schnurrbusch et al. reported the highest repeatability of the Spectralis OCT measurements in healthy eyes [5] for central retinal thickness and Comyn et al. in eyes affected with diabetic edema compared to other (SD-) OCT devices [3], Han and Jaffe [2] noticed up to 90.9 % of scans having at least one error in different macular disorders using the Spectralis™ for macular imaging [2], with the misidentification of the outer retina being the most common error of B-scans in 31.5% of B-scans. In this study similar rate of both: the overall occurrence of at least one segmentation error (88.4%), and the misidentification of the outer retina as the most frequent segmentation error (20.2%), was

observed. Indeed, the misplacement of the outer retinal boundary occurred especially in atrophic regions. This finding is supported by the high rate of errors noticed in the central and inner subfields. Furthermore it is reflected by the negative correlation between the error score and the TMV (as an indicator of atrophic processes). Typically, the misplacement of retinal boundaries led to false high measurements of retinal thickness and volume which affected these values significantly, and required careful review and manual correction (see supplemental figures 2).

The correct centration of the grid on the anatomical fovea was checked and adjusted, if necessary. As the grid is centered by the photographer according to institution's standard operating procedures, a possible initial misdisplacement of the grid by the software could not be evaluated.

### **Yearly change of TMV in different subgroups of patients**

A subgroup analysis of yearly change in TMV was conducted for different subgroups of patients and the results are presented together with the demographic information in supplemental table 3. This includes the results for patients with missing results from genetic testing and those with at least one confirmed mutation in the *ABCA4* gene. Furthermore, patients were categorized according to the classification of Fishman et al.[6] The estimated yearly change of retinal volume is similar between the patients with genetically confirmed mutations and the patients with missing genetic background testing. An analysis of subgroups of patients according to their clinical phenotype indicates the tendency of a higher yearly decline of TMV in group 3 versus group 2, and in group 2 than in group 1. However, the total number of patients in each of these subgroups is too small in order to provide a statistical significance.

### **Comparability of different scan densities**

In a subgroup of 20 eyes from 10 patients, SD-OCT imaging was repeated during the last visit with two different densities of scan lines (25 and 49 B-scans, respectively) of the 20° x 20° macular cube to establish intra-class correlation (ICC). First, the follow-up examination (examination A) was recorded prior to the acquisition of the second examination (examination

B) with the different scan density. After correction for segmentation errors, images of both examinations of the same eye and examination date were uploaded side-by-side on a screen. Care was taken not to change the ETDRS grid in the examination used for follow-up (examination A); the ETDRS-grid in examination B was re-placed by single grader (RWS) at the same spot as in examination A by direct comparison.

The ICC ( $\rho$ ) was estimated as the ratio of the between subject variance to the total variance:

$$\rho = \sigma_b^2 / (\sigma_b^2 + \sigma_w^2)$$

Where  $\sigma_b^2$  is the between subject variance and  $\sigma_w^2$  is the within subject variance. The SAS procedure NLMIX was used to perform the estimation. The ICC was calculated for 20 eyes of 10 patients for TMV, retinal volume in all nine ETDRS-subfields and the inner circle (inner four subfields) and outer circle (outer four subfields). Results are given in supplemental table 4. There is a good ICC of at least 0.943. That observation implies that in retrospective analyses these different scan densities could be used for analysis of data from different patients obtained with different scan densities.

### **Comparability between both eyes**

The correlation between both eyes during the first and the last visit was calculated. The correlation between eyes for total macular volume at first visit was 0.65 95% CI (0.36-0.94) and at the last visit, on average 1.3 years later, was 0.67 95% CI (0.39-0.94). Given the fact that STGD1 is a bilateral disease, this correlation may indicate that TMV may be considered as a suitable outcome measure for clinical trials. Moreover, the second eye can possibly serve as a control in local treatments that are applied to one eye only.

### **Genetic information**

Results from genetic testing through the EyeGene project (ClinicalTrials.gov Identifier: NCT00378742) were available for 14 patients: three patients showed one mutation, and eleven patients showed two mutations in the *ABCA4* gene. These are summarized in Supplemental table 5.

### **Supplemental text references:**

1. Yanni SE, Wang J, Cheng CS, et al. Normative reference ranges for the retinal nerve fiber layer, macula, and retinal layer thicknesses in children. *Am J Ophthalmol* 2013;155; 155(2):354: 354-60 e1.
2. Han IC, Jaffe GJ. Evaluation of artifacts associated with macular spectral-domain optical coherence tomography. *Ophthalmology* 2010;117(6):1177-89 e4.
3. Comyn O, Heng LZ, Ikeji F, et al. Repeatability of Spectralis OCT measurements of macular thickness and volume in diabetic macular edema. *Invest Ophthalmol Vis Sci* 2012;53(12):7754-9.
4. Cukras CA, Wong WT, Caruso R, et al. Centrifugal expansion of fundus autofluorescence patterns in Stargardt disease over time. *Arch Ophthalmol* 2012;130(2):171-9.
5. Wolf-Schnurrbusch UE, Ceklic L, Brinkmann CK, et al. Macular thickness measurements in healthy eyes using six different optical coherence tomography instruments. *Invest Ophthalmol Vis Sci* 2009;50(7):3432-7.
6. Fishman GA, Stone EM, Grover S, et al. Variation of clinical expression in patients with Stargardt dystrophy and sequence variations in the ABCR gene. *Arch Ophthalmol* 1999;117(4):504-10.

<b>Supplemental table 1: Frequency of different types of errors in the 6958 B-scans</b>	
<b>Type of error</b>	<b>Number of error</b>
Misidentification of inner retina (ILM)	138 (1.9%)
Misidentification of outer retina	1404 (20.2%)
Degraded scan image	10 (0.1%)
Out-of-register error (inner retina truncated because of a shifted scan)	118 (1.7%)
Inverted error in which the truncated superior and/or lateral edges of the retinal image were folded back onto the retina	2 (0.03%)
Incomplete segmentation error	503 (7.2%)
Complete segmentation error	22 (0.3%)
Cut-edge error	163 (2.3%)
Total number of errors	2360 (33.9%)

**Supplemental table 2:** Frequency of center involving algorithm errors in 190 reviewed macular cube scans is summarized \*.

<b>Stage of central involving error</b>	<b>Frequency</b>
Stage 0 (no error)	52.6 % (100/190)
Stage 1 (less than 1/3 of retinal thickness)	18.4% (35/190)
Stage 2 (More than 1/3, but less than 2/3 of retinal thickness)	21.1.% (40/190)
Stage 3 (more than 2/3 of retinal thickness)	34.7% (66/190)

*\* Center involving are errors that affect the central 1-mm-diameter area and/or the four inner subfields. Because some cubes show more than one segmentation error, the sum exceeds 100%.*

**Supplemental table 3:** Demographics of different subgroups of patients are given in respect to the status of genetic testing (presence or missing genetic confirmation of mutations in the ABCA4 gene) and for the phenotypical classification according to Fishman et al <sup>5</sup>. Based on different follow-up periods (3 to 40 months), the estimated yearly changes of total macular volume (TMV) are given for the patients with at least one follow-up visit for each subgroup.

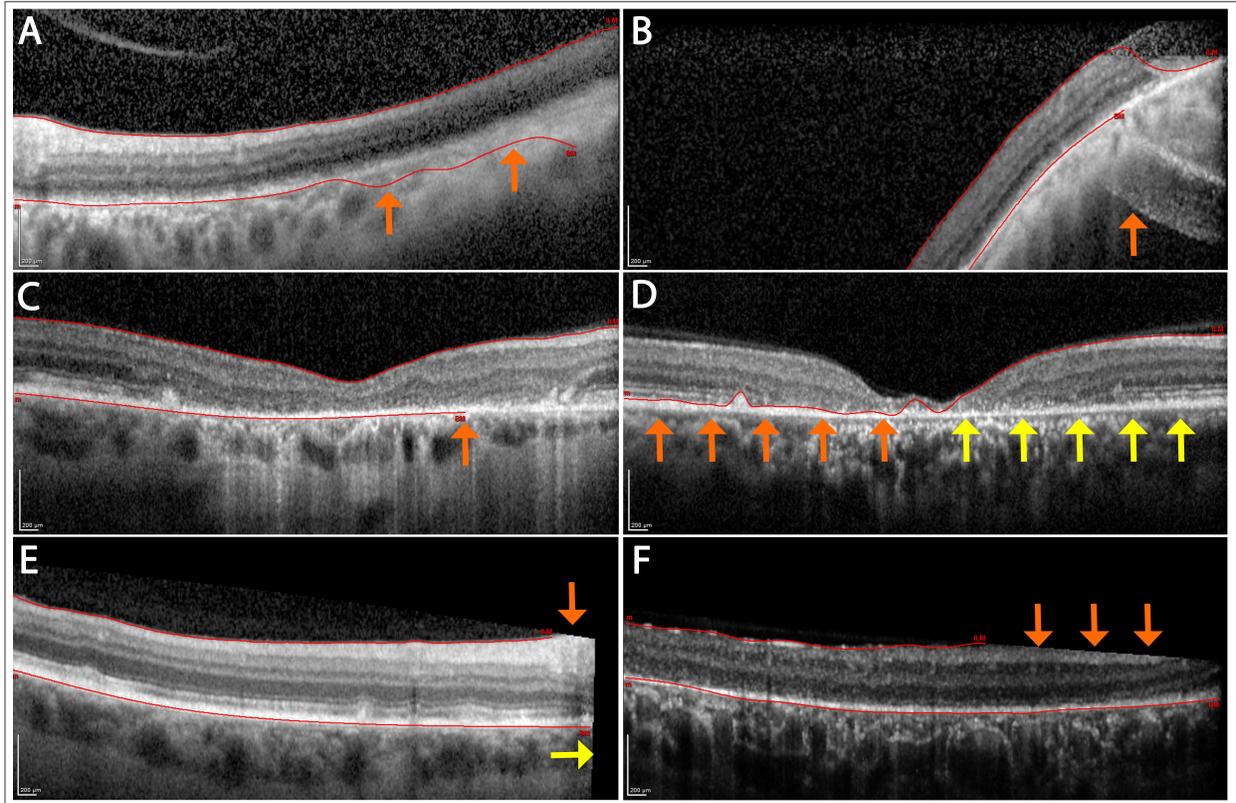
		Status of genetic testing of the 29 patients		Classification of the 29 patients according to Fishman et al.		
		Patients without results from genetic testing (n=15)	Patients with at least one confirmed mutation in the ABCA4 gene (n=14)	Group 1 (n=14)	Group 2 (n=10)	Group 3 (n=5)
Median age (range [years]) of onset of the disease		30.0 (12.0 – 53.)	26.0 (10.0 – 68.0)	21.5 (10.0 – 40.0)	33.5 (11.0 – 50.0)	50.0 (32.0 – 68.0)
Median age (range [years]) at first visit		41.0 (20.0 – 71.0)	34.5 (16.0 – 70.0)	29.5 (16.0 – 66.0)	41 (24.0 – 56.0)	66.0 (35.0 – 71.0)
Median (range) of best corrected visual acuity [logMAR] at first visit		0.35 (-1.0 – 0.9)	0.35 (-1.0 – 1.3)	0.4 (-0.1 – 0.9)	0.35 (-1.0 – 1.0)	0.25 (0 – 1.3)
Gender	male	9 (60.0 %)	7 (50.0 %)	6 (42.9%)	5 (50.0 %)	5 (100 %)
	female	6 (40.0 %)	7 (50.0 %)	8 (57.1%)	5 (50.0 %)	0 (0 %)
Race	Caucasian	8 (53.3 %)	11 (78.6 %)	11 (78.6 %)	6 (60.0 %)	2 (40.0 %)
	Afro-American	4 (13.3 %)	2 (14.3 %)	2 (14.3 %)	2 (20.0 %)	2 (40.0 %)
	Asian	2 (6.7 %)	0 (0 %)	0 (0 %)	2 (20.0 %)	0 (0 %)
	Indian	1 (3.3 %)	1 (7.1 %)	1 (7.1 %)	0 (0 %)	1 (20.0 %)
Yearly change of Total Macular volume [mm <sup>3</sup> ], estimated from the number of patients with at least two follow-up visits		-0.106 (n=10)	-0.120 (n=8)	-0.061 (n=7)	-0.122 (n=6)	-0.185 (n=5)

**Supplemental table 4:** Mean retinal thickness and volume in 20 eyes of patients during their last visit with two different image acquisition algorithms: one with a scan density of 25 lines and one with a scan density of 49 lines. In both algorithms, a scan of 20° x 20° was chosen, centered on the fovea.

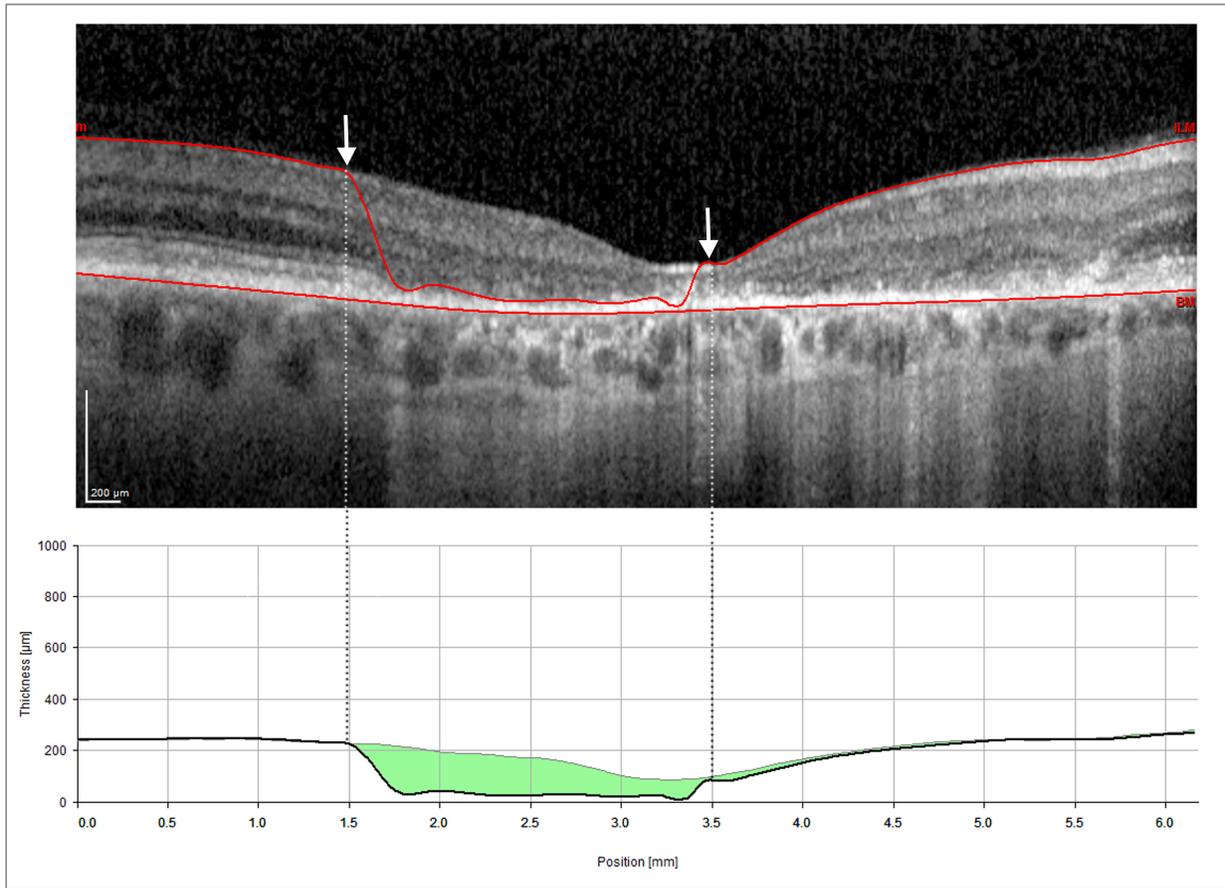
	25 scans Mean (SD)	49 scans Mean (SD)	ICC
Volume			
Total	6.396 (0.934)	6.404 (0.927)	0.999
Central	0.107 (0.039)	0.107 (0.040)	0.993
Superior inner circle	0.362 (0.064)	0.367 (0.058)	0.943
Nasal inner circle	0.336 (0.063)	0.336 (0.058)	0.966
Inferior inner circle	0.345 (0.051)	0.344 (0.051)	0.975
Temporal inner circle	0.323 (0.055)	0.327 (0.053)	0.968
Superior outer circle	1.359 (0.202)	1.357 (0.207)	0.998
Nasal outer circle	1.306 (0.225)	1.304 (0.228)	0.994
Inferior outer circle	1.255 (0.203)	1.251 (0.200)	0.995
Temporal outer circle	1.204 (0.185)	1.205 (0.176)	0.997

**Supplemental table 5:** Overview of confirmed mutations in the *ABCA4* gene in a subgroup of 14 patients. Genetic testing was performed in CLIA-certified laboratories.

ID	<i>ABCA4</i> Mutation Allele 1				<i>ABCA4</i> Mutation Allele 2			
	Exon	Position cDNA	Effect on protein	Nature of mutation	Exon	Position cDNA	Effect on protein	Nature of mutation
1	25	c.3608G>A	p.Gly1203Glu	missense	Intron 38	c.5460+4 A>G	IVS38+4 A>G	abnormal splicing
2	31	c.4577C>T	p.Thr1526Met	missense				
3	9	c.1222C>T	p.Arg408Stop	nonsense	42	c.5882G>A	p.Gly1961Glu	missense
4	1	c.1A>G	M1V	missense	42	c.5882G>A	p.Gly1961Glu	missense
5	1	c.1A>G	M1V	missense	42	c.5882G>A	p.Gly1961Glu	missense
6	Intron 38	c.5460+4A>G	IVS38+4A>G	abnormal splicing				
7	14	c.1957C>T	p.Arg653Cys	missense	24	c.3602T>G	p.Leu1201Arg	missense
8	19	c.2915C>A	pThr972Asn	missense	22	c.3322C>T	pARG1108Cys	missense
9	42	c.5882G>A	p.Gly1961Glu	missense	Intron 38	c.5461-10 T>C	IVS38-10 T>C	abnormal splicing
10	22	c.3322C>T	p.Arg1108Cys	missense	44	c.6079C>T	p.Leu2027Phe	missense
11	13	c.1927G>A	p.Val643Met	missense	22	c.3259G>A	p.Glu1087Lys	missense
12	22	c.3322C>T	p.Arg1108Cys	missense	44	c.6079C>T	p.Leu2027Phe	missense
13	46	c.6306C>A	p.Asp2102Glu	missense				
14	Intron 38	c.5460+4A>G	IVS38+4A>G	Abnormal splicing	42	c.5882G>A	p.Gly1961Glu	missense

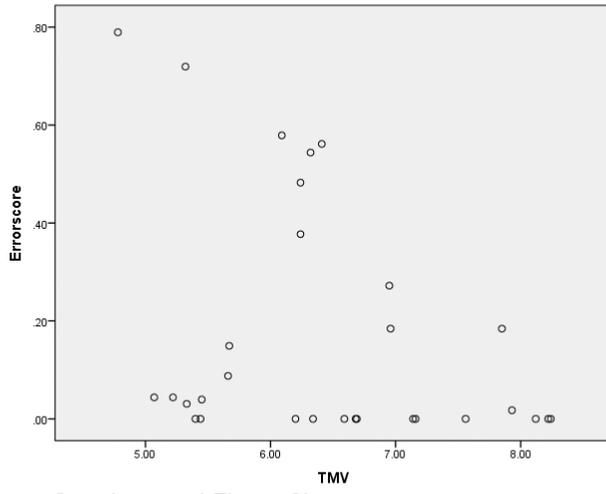


**Supplemental figures 1:** Examples of evaluated software errors are presented. **A:** Degraded scan image resulting in a misidentification of the outer retina and incomplete segmentation error (orange errors). **B:** Out-of register and inverted error, in which the lateral edges of the retina are folded back (orange arrow). Those errors were only noticed in two B-scans and didn't affect the measurement because they both happened at the edge of a macular cube scan. **C:** Incomplete segmentation error: the software stops the identification of the outer retina (orange error). **D:** Complete segmentation error: the software doesn't place any segmentation line at the outer retinal border (yellow arrows). Furthermore, the inner retinal boundary isn't recognized on the left half of the image and the according red segmentation line is misplaced back on the outer retinal boundary (orange arrows). **E:** Out-of-register (orange arrow) and cut-edge (yellow error) artifact. **F:** Out-of-register artifact (orange errors).

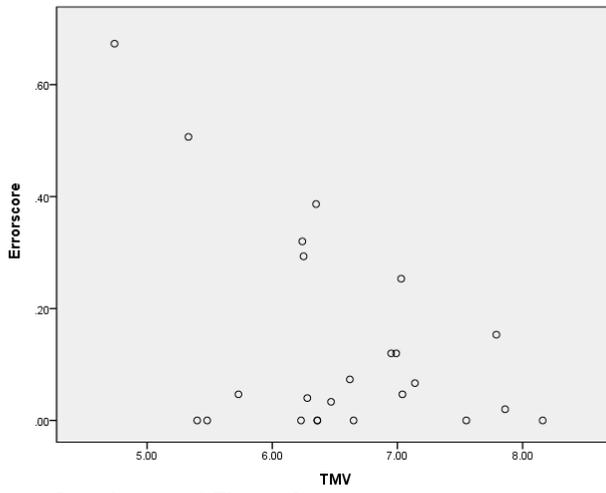


**Supplemental figure 2:** Example of a software error (misidentification of the inner retina) to illustrate the calculation of the error score. The horizontal length of each error in each line was registered. In this example, the horizontal length of the error measures 2 mm (between dotted lines), the total length of the single B-scan was 6 mm, respectively. The error score was calculated as the sum of error lengths/sum of total lines lengths of each macular cube. In errors affecting the central 1-mm-diameter area and four inner subfields, also the depth of each error was recorded and graded as stage 0 - no error; stage 1 – an error of one third or less of retinal thickness (mild error); stage 2 - an error more than one third, but less than two thirds of retinal thickness (moderate error); and stage 3 - more than two thirds of retinal thickness (severe error). This example shows a stage 3 error (error affecting more than two thirds of retinal thickness).

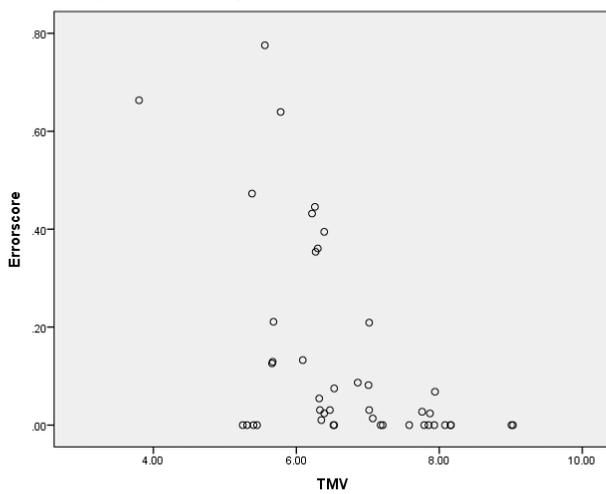
Supplemental Figure 3a



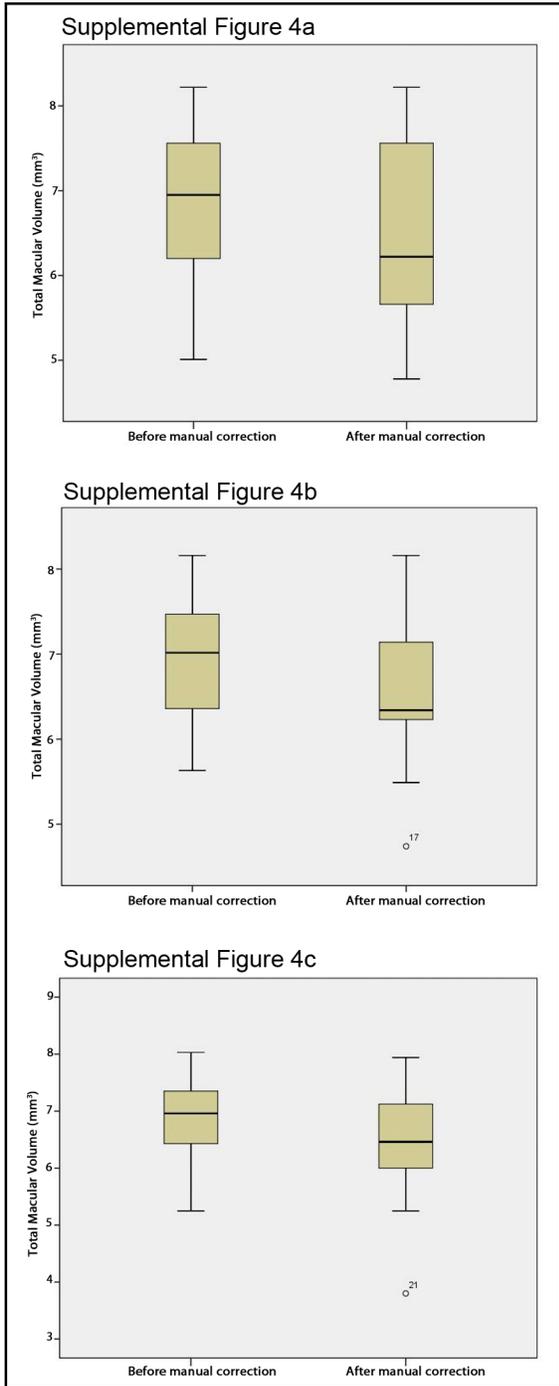
Supplemental Figure 3b



Supplemental Figure 3c



**Supplemental figures 3:** Shown are the correlations ( $r_s$ ) between error score and total macular volume (TMV, [mm<sup>3</sup>]) for the last visit of each patient for three different B-scan densities (supplemental figure 1a: 19 single B-scans ( $r_s = -0.413$ ,  $p = 0.023$ ), supplemental figure 1b: 25 single B-scans ( $r_s = -0.188$ ,  $p = 0.378$ ), and supplemental figure 1c: 49 single B-scans ( $r_s = -0.513$ ,  $p = 0.000$ ). The error score was calculated as: sum of all horizontal error lengths / sum of total line horizontal lengths of each macular cube.



**Supplemental figures 4:** Box-plots showing median and 25% quartiles of total macular volume (TMV) before and after manual correction of segmentation errors scans with three different scan-densities (18 macular cube scans with a density of 19 B-scans (supplemental figure 2a), 18 macular cube scans with a density of 25 single B-scans (supplemental figure 2a) and 32

macular cube scans with a density of 49 single B-scans (supplemental figure 2a)). There is a statistical significance before and after correction for all scan densities ( $p < 0.05$ ).