INTRODUCTION

Optical coherence tomography (OCT) is a non-invasive imaging modality for obtaining cross-sectional images of the retina in vivo. It is analogous to ultrasound, but instead of using the echoes created by acoustic waves, it uses light reflections to acquire information about structures. The ability of OCT to quantify retinal nerve fibre layer (RNFL) and macular thickness provides an objective method for monitoring axonal injury and serves as a useful outcome measure in clinical trials of optic nerve disorders.1–7

The University of California (UC) Davis Reading Center evaluated 19,961 scans from 981 subjects in two multiple sclerosis therapeutic trials with the aim of determining the influence of optical coherence tomography quality control procedures on error rates. There was no optical coherence tomography technician certification in Trial 1, and technicians had very limited monitoring and feedback during the trial in view of the fact that data were received retrospectively. However, technicians were certified in Trial 2 and submitted data in accordance with the protocol. Trial 2 scans had higher signal strengths, fewer errors, and more useable data than the scans in Trial 1. Thus, certified technicians and prompt transmission of data for ongoing quality control monitoring provided higher quality data in multiple sclerosis trials.

KEYWORDS multiple sclerosis, neuro-ophthalmology, OCT, quality control, reading centre

ABSTRACT

The University of California (UC) Davis Reading Center evaluated 19,961 scans from 981 subjects in two multiple sclerosis therapeutic trials with the aim of determining the influence of optical coherence tomography quality control procedures on error rates. There was no optical coherence tomography technician certification in Trial 1, and technicians had very limited monitoring and feedback during the trial in view of the fact that data were received retrospectively. However, technicians were certified in Trial 2 and submitted data in accordance with the protocol. Trial 2 scans had higher signal strengths, fewer errors, and more useable data than the scans in Trial 1. Thus, certified technicians and prompt transmission of data for ongoing quality control monitoring provided higher quality data in multiple sclerosis trials.

KEYWORDS multiple sclerosis, neuro-ophthalmology, OCT, quality control, reading centre
beginning of the trial, whereas Trial 1 technicians had essentially no monitoring or feedback during the trial.

In this article, we describe the QC procedures for OCT scans and report on the comparison between the qualities of the OCT data in these two multi-centre MS therapeutic trials.

**MATERIALS AND METHODS**

The UC Davis OCT Reading Center evaluated 19,961 OCT scans from 981 subjects enrolled in the two MS therapeutic trials. Trial 1 involved 939 relapsing-remitting MS patients from 96 sites that submitted 18,733 OCT scans. Trial 2 involved 40 secondary progressive MS patients from 10 sites that submitted 2178 OCT scans. At the start of these trials, high-resolution OCT techniques were not widely available. Thus, all scans were supposed to be performed using Stratus TM time-domain OCT (Carl Zeiss Meditec, Inc., Dublin, CA) and were supposed to measure the retinal area encompassing the macula and the thickness of the retinal nerve fibre layer around the optic disc. Acceptable reproducibility has been reported by experienced centres using the Stratus OCT.12 It has historically been used to quantify RNFL thinning in patients with multiple sclerosis (MS) and/or optic neuritis under idealized conditions.1,3,4,17–21 The Stratus OCT with Retinal Nerve Fibre Layer (RNFL) Normative Database offers clinicians a quantitative tool for the comparison of the retinal nerve fibre layer in the human retina to a database of known normal subjects.22,23

Trial 1 required a total of 4 scans per visit: Fast Macular Thickness Map (2) and Fast RNFL Thickness (2). Trial 2 required a total of 18 scans per visit; Fast Macular Thickness Map (6), Fast RNFL Thickness (6), and RNFL Thickness (6).

**TECHNICIAN CERTIFICATION**

The Reading Centre certified technicians for OCT in Trial 2. Certification required passing a 10-question telephone examination to demonstrate familiarity with the data collection requirements and export of sample scans required for the study. Technician certification in some cases required repeated testing after feedback from the Reading Center. Following certification, the Reading Center continually monitored the data and provided feedback to the coordinators at the clinical sites. In Trial 1, the coordinators submitted OCT data based upon the Reading Center’s written protocol. However, there was no technician certification in Trial 1. Ongoing monitoring and feedback from the Reading Center was limited as retrospective data were received in batches. Any feedback given to the technicians was given after the scans had already been acquired and received by the Reading Center. The Reading Center received very few follow-up data after the QC information on retrospective data had been provided thus improvement based on this follow-up data could not be adequately measured.

**Quality Control Assessment**

The Reading Center’s central readers (first and last authors, J.K. and J.W.) reviewed and evaluated all 19,961 OCT scans for quality control (QC) measures.

**Test Parameter and Shipping Error Assessment**

The quality control system addressed three areas of clinic performance: test parameter errors, shipping errors, and scan quality (proper alignment and usability). Test parameter errors included signal strength <7, extra scans performed, wrong scan used, scans performed on the wrong instrument, and scans requiring redraws (due to segmentation artifacts). Shipping errors included exported data missing and missing scans.

**Unusable Scans (Objective Method)**

A scan was determined to be unusable if it was missing, performed as a wrong scan, or performed on the wrong instrument.

**Scan Quality for Proper Alignment and Usability (Masked Evaluation)**

Because Trial 1 did not have technician certification at any time, we decided to perform a masked analysis on scan quality to remove any bias that may have occurred between reviewing the two trials. The two readers (J.K. and J.W.) evaluated and compared a subgroup analysis of 1200 randomly chosen scans (Trial 1 = 725; Trial 2 = 475) for proper alignment and usability in a masked fashion. A scan was determined to be unusable if it was performed with a low signal strength (<7), contained an artifact, or was not properly aligned or both. Furthermore, each scan was individually reviewed by the Reading Center, and parameters were entered in a 4th dimension database software program. Data were assigned a 100-point value and penalized according to any errors or protocol deviations.

**RESULTS**

**Test Parameter Errors**

As shown in Table 1, more protocol deviations occurred in Trial 1 than in Trial 2. A total of 23.2% (4,353/18,733) of the scans in Trial 1 had a signal strength <7, whereas 3.7% (81/2,178) had a signal strength <7 in Trial 2 (Figure 1). A substantial number of extra scans were performed in
TABLE 1 Test parameter and shipping errors quality control summary.

<table>
<thead>
<tr>
<th>Test parameter</th>
<th>Trial 1 ( N = 939 ) subjects 96 sites</th>
<th>%</th>
<th>Trial 2 ( N = 40 ) subjects 10 sites</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total scans</td>
<td>18,733</td>
<td></td>
<td>2178</td>
<td></td>
</tr>
<tr>
<td>Errors:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Signal strength &lt;7</td>
<td>4353</td>
<td>23.2%</td>
<td>81</td>
<td>3.7%</td>
</tr>
<tr>
<td>Manual entry (exported data missing)</td>
<td>4114</td>
<td>22.0%</td>
<td>0</td>
<td>0%</td>
</tr>
<tr>
<td>Extra scans performed</td>
<td>2537</td>
<td>13.5%</td>
<td>29</td>
<td>1.3%</td>
</tr>
<tr>
<td>Missing scans</td>
<td>475</td>
<td>2.5%</td>
<td>34</td>
<td>1.6%</td>
</tr>
<tr>
<td>Wrong scans</td>
<td>448</td>
<td>2.4%</td>
<td>42</td>
<td>1.9%</td>
</tr>
<tr>
<td>Wrong instrument used</td>
<td>269</td>
<td>1.4%</td>
<td>0</td>
<td>0%</td>
</tr>
<tr>
<td>Topcon scans</td>
<td>42</td>
<td>0.2%</td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td>Cirrus scans</td>
<td>218</td>
<td>1.2%</td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td>Heidelberg</td>
<td>9</td>
<td>&lt;1%</td>
<td>n/a</td>
<td></td>
</tr>
<tr>
<td>Redraws</td>
<td>126</td>
<td>&lt;1%</td>
<td>0</td>
<td>0%</td>
</tr>
</tbody>
</table>

STRATUS OCT Review Software
RETINAL MAP ANALYSIS Report - 5.0.1 (0376)

FIGURE 1 An arrow indicates a low signal strength in a fast macular thickness map of the right eye.
Trial 1 compared with Trial 2. A total of 2537 (13.5%) extra scans were generated in Trial 1, compared with 29 (1.3%) in Trial 2. The wrong instrument was used more frequently in Trial 1, with 269 (1.4%) scans performed on the Cirrus (218), Topcon (42), or Heidelberg (9) instruments. Trial 2 had no scans performed on the wrong instrument. Scans requiring redraw of the segmentation were more prevalent in Trial 1 than in Trial 2, with 126 scans in Trial 1 versus 0 scans in Trial 2.

**Shipping Errors**

Both trials required that OCT data be downloaded to a USB drive and sent to the *Reading Center* for processing. Trial 1 had 22.0% (4,114/18,733) of the scans missing the exported data whereas Trial 2 had 0 scans with the same error. Without electronic data transfer, the *Reading Center* was responsible for manually entering the scans in the database, creating another potential source of error. Trial 1 also had 475 (2.5%) of the scans missing (scans not performed) whereas Trial 2 had 34 (1.6%).

**Unusable Scans (Objective Method)**

There were 475 (2.5%) missing scans, 448 (2.4%) scans performed using the wrong scan, and 269 (1.4%) scans performed on the wrong instrument in Trial 1 for a total of 1,192 (6.4%) unusable scans. In Trial 2, however, there...
were 34 (1.6%) missing scans, 42 (1.9%) wrong scans, and 0 (0%) scans performed on the wrong instrument for a total of 76 (3.5%) unusable scans.

**Scan Quality for Proper Alignment and Usability (Masked Evaluation)**

Proper centering of the macula and optic disc are essential for accurate RNFL thickness measurements. Slight misalignment can cause large deviations from the correct data values and generate erroneous data as shown in Figures 2 and 3. From the 1200 randomly chosen scans, 59.3% (429/725) of the scans were properly centered in Trial 1, whereas 74.7% (355/475) of the scans in Trial 2 were properly centered (Table 2). The readers determined that 21% (150/725) of the scans in Trial 1 were not useable, whereas 6% (29/475) of the scans were not useable in Trial 2. Scans not properly centered have erroneous values and will most likely be rendered unusable for data analyses. Figure 4 demonstrates RNFL scans that were rendered unusable because the macula was imaged instead of the optic nerve.

**DISCUSSION**

The ability of OCT to quantify RNFL and macular thickness allows an objective method for monitoring...
Neuro-ophthalmology

John L. Keltner et al.

axonal injury and potentially can serve as a useful outcome measure in clinical trials of MS. In MS clinical trials, sequential OCT imaging may serve as a potential surrogate structural marker for, or complement to, the more time-consuming and expensive magnetic resonance imaging (MRI). This non-invasive technique thus may be useful in directing clinical response to pharmacotherapy.\textsuperscript{1,6,16,21} Although OCT will not replace MRI as an objective measure of MS progression, it may be a useful adjunct that provides additional information as a potential marker of neuronal pathology.

Patients with MS and a history of unilateral optic neuritis have RNFL thinning not only in affected eyes, but also in the unaffected fellow eyes as demonstrated by TD-OCT.\textsuperscript{2,4,18,26} MS patients with no history of acute optic neuritis have shown decreased RNFL thickness compared with eyes of healthy subjects, as measured by Stratus TD-OCT. This decrease has been found to correlate well with low-contrast letter acuity and contrast sensitivity in such patients.\textsuperscript{4,18} These findings are consistent with RNFL thinning in MS patients occurring on a chronic basis and not exclusively from acute optic neuritis, further warranting the use of OCT to follow disease progression and response to therapy.

Measurements of optic nerve atrophy in patients with MS and optic neuritis with MRI have recently been...
correlated with optic nerve thinning as measured by TD-OCT, further validating RNFL measurement as a potential surrogate structural marker for central nervous system imaging in clinical and research investigations in MS. As retinal ganglion cells make up about one-third of the total macula thickness, attention has also been focused on following macular thickness reductions in demyelinating disease. An association between optic nerve RNFL thinning and macular volume reduction in patients with optic neuritis with or without MS has been reported with TD-OCT.

The present paper shows the importance of QC in obtaining high-quality OCT data from MS clinical trials. The difference between the size of Trial 1 (939 subjects) and the size of Trial 2 (40 subjects) in this study is significant; thus it is important to note that larger studies are logistically going to be much more complicated than smaller ones. As shown in Trial 2, technicians who are trained and certified in a standard fashion generally produce reliable, high-quality OCT data by reducing signal strength errors, extra/unneeded scans, manual entry, missing scans, use of wrong instruments, and redraws of the segmentations. The dramatic differences between the quality of the data in these two MS clinical trials is well demonstrated in Tables 1 and 2. We believe these results are not surprising since there was no OCT technician certification in Trial 1, and technicians had very limited monitoring and feedback during the trial since data was received retrospectively.

To obtain high-quality data, OCT reading centres should be involved at the beginning stages of planning for any clinical trial in MS. Prompt transmission of OCT data to the reading centre for ongoing and intensive QC monitoring and rapid feedback to technicians helps to reduce the frequency of incorrect testing parameter and shipping errors. Many of the issues encountered in Trial 1 we believe could have been avoided had our reading centre been involved at the beginning of the MS clinical trial. Signal strength and the presence of artifacts have been shown to affect RNFL thickness measurements using Stratus OCT.

In summary, the quantitative data and examples in Figures 1 to 4 demonstrate the utility of an OCT reading centre in MS clinical trials using OCT technology to monitor MS patients with RNFL or macula volume changes. To optimise data quality, it seems essential for an OCT reading centre to be involved from the earliest planning stages of the trial to ensure technicians are properly certified for OCT testing. Even after certification, it is important to provide continuous monitoring with rapid and constructive feedback upon submission of the data.

**ACKNOWLEDGEMENT**

These studies are based on industry-supported trials and are supported by Research to Prevent Blindness (RPB).

**Declaration of interest:** The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

**Note:** Figures 1-4 of this article are available in colour online at www.informahealthcare.com/oph.

**REFERENCES**


