Repeatability and Reproducibility of Corneal Thickness Measurements by Optical Coherence Tomography

Sarah Muscat, Nicola McKay, Stuart Parks, Ewan Kemp, and David Keating

PURPOSE. To assess the repeatability and interoperator and intersession reproducibility of central corneal thickness (CCT) measurements made by a commercially available optical coherence tomography (OCT) system.

METHODS. Intersession and interoperator reproducibility and repeatability were defined according to the guidelines of the British Standards Institution and examined in a control group of 14 normal subjects. An in-house computer program was used to evaluate central corneal thickness from these scans.

RESULTS. The coefficient of interoperator reproducibility was 0.18%, whereas that for intersession reproducibility was 1.11%. Wilcoxon analysis (5% level of significance) showed that there was no statistically significant difference between scans acquired during different sessions or by different operators. Coefficients of repeatability were all less than 3%. The average CCT was $526 \pm 28 \mu m$ (SD) and the range of normal CCT between 5th and 95th percentiles was 498 to 576 \mu m.

CONCLUSIONS. Although the commercially available OCT scanner was designed for retinal imaging, with a few minor modifications, the system may be used to image the anterior segment. Previous studies have shown that OCT measurements correlate well with those from conventional techniques, and it has the added advantage of being a noncontact technique. This study further demonstrates that the OCT measurements show a high degree of repeatability and reproducibility. Thus, OCT is emerging as a promising tool for evaluation of CCT in the clinical setting. (Invest Ophthalmol Vis Sci. 2002;43:1791–1795)

Measurement of central corneal thickness (CCT) is performed for both diagnostic and therapeutic purposes. The increasing popularity of refractive surgical procedures has made accurate measurement of CCT a matter of paramount importance. Moreover, analysis of corneal thickness in contact lens wearers is essential for monitoring any changes in the cornea. Knowledge of CCT is also necessary for accurate determination of intraocular pressure, because tonometer readings are dependent on corneal thickness to a certain degree. This has important implications in the management of glaucoma.

There are various ways of measuring corneal thickness. The most commonly used clinical methods are optical and ultrasound pachymetry, and there are a number of published studies that address the reliability and repeatability of these measurement techniques.\textsuperscript{1–9} Optical coherence tomography (OCT) is a relatively new imaging technique that makes it possible to obtain cross-sectional images of the eye in vivo and noninvasively. To date, this technique has mostly been applied to imaging of retinal structure but can also be used to obtain cross-sectional images of the cornea.

One research group has investigated the use of OCT in corneal imaging, and it has been shown that CCT measurements made with the OCT system correlate very well with measurements made with other techniques.\textsuperscript{10} OCT has an added advantage over techniques, such as ultrasound pachymetry, in that no contact with the eye itself is necessary, and thus it can be used soon after corneal surgery.

The commercially available OCT system (Humphrey-Zeiss Medical Systems, San Leandro, CA) has been designed to image the retina rather than the anterior eye; however, images of the cornea are quite easy to obtain. Several studies have shown that the system yields highly repeatable and reproducible measurements of retinal thickness.\textsuperscript{11–15} Because OCT is also emerging as a promising tool for evaluation of the cornea, it is important to quantify these parameters for corneal thickness measurements as well, especially because the system is not really designed for imaging the anterior eye. A few studies have examined this matter to a limited extent\textsuperscript{10,16}; however, a more rigorous approach is needed. Thus, the purpose of this study was to perform corneal thickness measurements with OCT in a control group of normal volunteers and to determine intersession reproducibility, interoperator reproducibility, and repeatability of these measurements, according to the definitions set by the British Standards Institution and the International Standards Organization.\textsuperscript{17,18}

MATERIALS AND METHODS

The scanner used in this study is a commercially available OCT scanner (Humphrey-Zeiss Medical Systems). Scanning is performed using a superluminescent diode operating in the near infra-red, with a wavelength of 850 nm and maximum power of 750 \mu W. Each B-scan consists of 100 A-scans, regardless of the length of the scan line, and images are displayed as a pseudocolor plot on which different colors represent differences in the reflective properties of the corneal tissue. Although this equipment is designed for posterior, rather than anterior, eye imaging, corneal scans are easily obtained with very minor modifications to the system. To focus the laser beam onto the cornea for corneal scanning, we found that patients had to be positioned slightly farther away from the machine head than in the setup for retinal imaging. We therefore attached a small cushion to the forehead rest to move the patients’ eyes away from the machine by a few centimeters and still ensure comfortable and steady positioning of the head.

Subjects

Fourteen volunteers (seven men and seven women), ranging in age from 21 to 58 years (average age, 33.4 years) participated in this study. The study was conducted according to the tenets of the Declaration of Helsinki, and volunteers gave informed consent after the nature and intent of the study had been fully explained to them. The exclusion criteria were any ocular abnormalities, history of eye disease, prior refractive surgery, and contact lens wear. All scanning was performed in the undilated left eye. During scanning, subjects were asked to fix on...
the external fixation light with the right eye. The external fixation light was positioned so that the subject was looking straight ahead.

**Scanning**

Our definitions of repeatability and reproducibility were based on the definitions adopted by the British Standards Institution. Under repeatability conditions, independent test results are obtained with the same method, on the same subject, by the same operator, and on the same set of equipment with the shortest time lapse possible between successive sets of readings. We investigated repeatability by obtaining 10 horizontal and 10 vertical scans, ensuring that each scan bisected the pupil. All scanning was performed by the same operator. The time elapsed between successive scans corresponded to the time taken to save the previous scan and adjust the position of the new scan if necessary and never amounted to more than a few seconds. Repeatability was investigated in three subjects. The scan length on the OCT system was set to 2 mm, using the default axial length of 24.46 mm. However, because the system is designed for retinal imaging, this is not the true length of the scan line on the cornea. The scan length was calibrated by projecting the scan line onto a ruler placed at the approximate position of the cornea, and the true scan length was found to be approximately 2.5 mm.

Under reproducibility conditions, sets of readings are obtained using the same method but with one variation in the experimental setup. In this study, we investigated both intersession and interoperator reproducibility. In both studies, sets of six radial scans, centered on the center of the pupil were obtained by using the radial-lines-scanning option. The scan length set on the machine was 2 mm (radial scan length set to 1 mm); thus, the true calibrated scan length was approximtely 2.5 mm. Intersession reproducibility was investigated by acquiring readings on two separate days by a single operator. All scanning was performed during a specific period—11:00 AM to 2:00 PM—to minimize the effect of diurnal variations in corneal thickness. Studies have shown that CCT is most stable during this period.

To investigate interoperator reproducibility, two operators each obtained a set of six scans on every subject during the same scanning session. The subject was asked to sit back and relax for a few minutes after the first set of scans while the first operator exited the OCT system. The second operator then repositioned the subject and started the system once again before obtaining the second set of scans. Both intersession and interoperator reproducibility were investigated for a control group of 14 subjects.

**Analysis**

A typical OCT scan through the center of the cornea is shown in Figure 1A. This scan is composed of 100 A-scans. The A-scan at the location marked by the white arrow in Figure 1A is shown in Figure 1B. The anterior and posterior surfaces of the cornea are easily identified as the first and last spikes in the image, and corneal thickness is measured by calculating the longitudinal displacement between the positions of these two maxima. The raw OCT data were exported and analyzed by a custom-built computer program designed to identify these two peaks, place cursors on them and calculate the distance between them. After running this program, each A-scan was checked to ensure that the program had correctly identified the peaks. In the few cases in which the system had failed to identify the peaks correctly, the cursors were manually placed in the correct position, thus ensuring that all thickness measurements were correct.

The OCT system is designed to image a concave structure, the retina, and has in-built correction factors to compensate for the curvature of the retina. Thus, when using the system to image a convex structure such as the cornea, the system exaggerates the curvature. Moreover, when the subject is looking straight ahead, the OCT beam is perpendicular to the surface of the cornea only at the center, and measurements made off-center therefore overestimate the corneal thickness. In view of this, we did not attempt to quantify peripheral corneal thickness, and limited ourselves to measuring only CCT. For each scan, we identified the apical A-scan, as shown in Figure 1A and defined this as the central A-scan. We then looked at 10 A-scans to the left of the central one and 10 to the right and quantified the corneal thickness from each of these 21 A-scans. Because the approximate length of the whole scan line was 2.5 mm and the scan consisted of 100 equally spaced A-scans, these central 21 scans spanned a length of 0.5 mm across the cornea. We defined this 0.5-mm section as the central cornea and assumed that the corneal surface was relatively flat and perpendicular to the OCT beam in this region.

In the reproducibility study we acquired six scans arranged along the diameters of a circle centered on the center of the pupil (using the radial-lines-scanning option) in each subject during each session. In this scanning mode the OCT system acquires a set of six scans of equal length arranged along six diameters of a circle of predetermined size. Each of these scans takes 1 second to acquire, and the system takes approximately 2 seconds to save each scan. Thus, assuming that the patient does not move between scans and that no repositioning of the scan lines is therefore required, this set of six scans should take approximately 30 seconds to acquire. If any patient movement was detected between scans, the whole set was discarded and a new set was acquired. For each of these scans, we selected the central 21 A-scans, as explained, and made corneal thickness measurements at each of these points. This yielded 120 corneal thickness measurements in each patient, arranged at different points along the radial scan lines and 6 repeated measurements at the center, at the point where the six scans intersect (Fig. 2).

As suggested by Bland and Altman who based their definitions on the recommendations of the British Standards Institution, the coefficient of reproducibility was defined as the SD of the differences between pairs of measurements obtained during different sessions or by different operators, divided by the average of the means of each pair of readings. For each patient in the reproducibility study we first randomly selected one of the six of corneal thickness measures at the intersection point of the six scans obtained during one session and then calculated the overall average corneal thickness from this value.
together with the other 120 thickness measures obtained along the radial lines. These results were used to calculate the coefficients of intersession and interoperator reproducibility. Graphs of differences against means, as suggested by Bland and Altman,\textsuperscript{22} were plotted for both the interoperator and the intersession studies. In both cases, the Wilcoxon matched-pairs test (5% significance level) was also used to establish whether there was any statistically significant difference between measurements made during different sessions or by different operators.

To establish the characteristics of OCT corneal scans in the normal population, we first computed a single corneal thickness for each subject, representing the average corneal thickness over the area of central cornea under consideration. These measures were then averaged over all the subjects. We also calculated separate means for the men and women and investigated whether there were any gender-related statistically significant differences in corneal thickness. We corrected thicknesses for refractive index, because the refractive index adopted by the system (1.38) and that of the cornea (1.376) are very similar.

The coefficient of repeatability obtained from the repeated administration of the test under identical conditions was defined as the SD of the difference from the mean of these repeat measurements divided by the average response. Coefficients of repeatability were calculated from the 10 consecutive vertical and horizontal scans in each of the three subjects participating in the repeatability study.

\section*{RESULTS}

\subsection*{Central Corneal Thickness}

The overall mean \(\text{CCT} = 526 \pm 28 \ \mu m \) (SD). Mean \(\text{CCT} \) for males was \(528 \pm 25 \ \mu m\), whereas in the women it was \(524 \pm 33 \ \mu m\). The median and the 5th and 95th percentiles of \(\text{CCT}\) were 519, 498, and 576 \(\mu m\), respectively. Wilcoxon analysis (5% level of significance) showed that there was no statistically significant difference between the male and female groups. Intrapatient variability (the mean intrapatient SD) was found to be 3.7 \(\mu m\).

\subsection*{Interoperator Reproducibility}

In each subject, the overall average corneal thickness was calculated by operator 1 and operator 2. The coefficient of reproducibility was then computed from these data and was found to be 0.18\%. The intraclass correlation coefficient (ICC) for interoperator reproducibility was 0.998. Figure 3 shows the graph of differences against means of the average corneal thickness. All data were within 2 SDs of the mean. According to the definitions of the British Standards Institution, this indicates reproducibility in \(\text{CCT}\) measurements acquired by two operators during the same session. From this graph, the 95\% limits of agreement (LoA), defined as mean interoperator difference in thickness \(\pm (1.96 \times \text{SD of differences})\), was -3 to +4 \(\mu m\). The Wilcoxon paired measurement test (5\% significance level) was also performed on the thickness measures. No statistically significant differences were found between the two sets of data. In a study involving 14 participants, the probability is 90\% that the study will detect a difference in measurements made by different operators if the true difference is 1.75 \(\mu m\) (two-sided 5\% significance level).

\subsection*{Intersession Reproducibility}

Intersession reproducibility was investigated in a similar way. The coefficient of intersession reproducibility was found to be 1.11\%. From the graph of differences against mean for the intersession data (Fig. 4), it can be seen that 100\% of the values fall within 2 SDs of the mean. The 95\% LoA are -16 to +7 \(\mu m\). The ICC for intersession reproducibility was 0.979. Wilcoxon analysis showed that there was no statistically significant difference between the two groups. In a study involving 14 participants, the probability is 90\% that the study will detect a difference in measurements acquired by the same operator but during different sessions if the true difference between sessions is 5.45 \(\mu m\) (two-sided 5\% significance level).

\begin{figure}
\centering
\includegraphics[width=\textwidth]{fig3}
\caption{Graph of data from interoperator reproducibility study. Mean corneal thickness in each subject was plotted against difference in corneal thickness, as measured by operators 1 and 2. All measurements \((n = 14)\) were within 2 SDs of the mean.}
\end{figure}

\begin{figure}
\centering
\includegraphics[width=\textwidth]{fig4}
\caption{Graph of data from interoperator reproducibility study. Mean corneal thickness in each subject was plotted against difference in corneal thickness, as measured by operators 1 and 2. All measurements \((n = 14)\) were within 2 SDs of the mean.}
\end{figure}
Repeatability

For each of the three subjects in the repeatability study, we computed the coefficient of repeatability for the 10 horizontal and 10 vertical scans. The average of the horizontal sets was 1.76%, whereas that of the vertical scans was 2.32%. The individual results are displayed in Table 1.

DISCUSSION

The potential of OCT for corneal imaging was first shown by Izatt et al.23 with the prototype OCT system that was later developed into the commercially available system. More recently, Drexler et al.24 have shown images of the cornea with even higher resolution obtained with a prototype ultrahigh-resolution system. With a few very minor modifications, a commercially available system (OCT; Humphrey Zeiss) originally designed for retinal imaging can also be used to image the cornea. Several studies give examples of various situations in which OCT was useful in the evaluation of corneal conditions, especially after refractive surgery.16,25–28 In addition, OCT may be used to make measurements of corneal thickness and has an advantage over more established methods of measuring corneal thickness, in that it is a noncontact and noninvasive technique. Corneal thickness measured with OCT has been found to correlate very well with measurements from ultrasound pachymetry, as shown by Bechmann et al.10 who found correlation to be higher with this new high-resolution system. Marsich and Bullimore1 investigated the performance of three devices for determining corneal thickness and found that the Orbscan slit lamp pachymetry system (Orbscan Topography System I; Bausch & Lomb, Inc., Salt Lake City, UT) showed the best repeatability (what we have termed interobserver reproducibility) with 95% LoA of −10 to +17 μm. We obtained 95% LoA of −16 to +7 μm from the intersession data. This indicates that the performance of OCT is comparable to that of established methods of evaluating corneal thickness.

Our average CCT was 526 ± 28 μm, almost identical with the 530 ± 32 μm obtained by Bechmann et al.10,29 in normal subjects. This group found that the interpatient variability with OCT was 4.90 μm, whereas interpatient variability was 32 μm, similar to our findings. A study by Radhakrishnan et al.30 describes a higher-resolution OCT system designed specifically for corneal imaging. This study quotes a corneal thickness of 488 μm measured in a single volunteer. This thickness is lower than that obtained in our study. Whether this is because the measurement came from a single subject or because the commercial system overestimates corneal thickness in comparison with this new high-resolution system remains to be seen. Our study showed no gender-related differences in average corneal

Although the system used in this study (OCT; Humphrey-Zeiss) is known to yield repeatable and reproducible measurements of retinal and nerve fiber layer thickness,11–15 we thought it also necessary to rigorously assess its performance in measuring corneal thickness, especially because this commercial system is not actually designed to perform imaging of the anterior segment. In our study, we evaluated repeatability, intersession reproducibility, and interoperator reproducibility as suggested by Bland and Altman,22 who based their definitions on the standards set by the British Standards Institution.17,18

The coefficients of repeatability for vertical and horizontal scanning were less than 5% in all subjects in this part of the study. This indicates that during the same scanning session, corneal thickness measurements are repeatable, and there is therefore no need to acquire a large number of readings for a reliable estimate of CCT.

Coefficients of interoperator and intersession reproducibility were 0.18% and 1.11%. The observation that the intersession reproducibility is slightly higher than the interoperator reproducibility may be an indication of diurnal variation in corneal thickness between different sessions. Although we attempted to acquire the second set of scans at approximately the same time of day, but on a different day from the first, it appears that there was nevertheless a small amount of variation. Marsich and Bullimore1 investigated the performance of three devices for determining corneal thickness and found that the Orbscan slit lamp pachymetry system (Orbscan Topography System I; Bausch & Lomb, Inc., Salt Lake City, UT) showed the best repeatability (what we have termed interobserver reproducibility) with 95% LoA of −10 to +17 μm. We obtained 95% LoA of −16 to +7 μm from the intersession data. This indicates that the performance of OCT is comparable to that of established methods of evaluating corneal thickness.

Our average CCT was 526 ± 28 μm, almost identical with the 530 ± 32 μm obtained by Bechmann et al.10,29 in normal subjects. This group found that the interpatient variability with OCT was 4.90 μm, whereas interpatient variability was 32 μm, similar to our findings. A study by Radhakrishnan et al.30 describes a higher-resolution OCT system designed specifically for corneal imaging. This study quotes a corneal thickness of 488 μm measured in a single volunteer. This thickness is lower than that obtained in our study. Whether this is because the measurement came from a single subject or because the commercial system overestimates corneal thickness in comparison with this new high-resolution system remains to be seen. Our study showed no gender-related differences in average corneal

### TABLE 1. Vertical and Horizontal Coefficients of Repeatability for the Three Subjects in the Repeatability Study

<table>
<thead>
<tr>
<th>Subjects</th>
<th>Horizontal Coefficient of Repeatability</th>
<th>Vertical Coefficient of Repeatability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject 1</td>
<td>1.63</td>
<td>2.08</td>
</tr>
<tr>
<td>Subject 2</td>
<td>1.89</td>
<td>1.94</td>
</tr>
<tr>
<td>Subject 3</td>
<td>1.76</td>
<td>2.94</td>
</tr>
</tbody>
</table>

Data are percentages.
Repeatability and Reproducibility of CCT Measured by OCT

thickness; however, our sample sizes (seven men and seven women) were rather small.

In our study we limited ourselves to assessing only CCT. Peripheral measurements would have been inaccurate for two reasons: First, at peripheral points the OCT beam would not have been perpendicular to the corneal surface, and measurements therefore would have had to be corrected for refraction at the air–cornnea interface; second, the radius of curvature seen on the corneal scans is overestimated, because the system is designed to image a concave rather than convex structure and therefore exaggerates the curvature of the cornea. CCT measurements, however, have been shown to be both repeatable and reproducible, and this indicates that the commercially available OCT system in its present state and with practically no modifications is a reliable and useful tool in assessing and monitoring corneal conditions.

References