Repeatability of macular thickness measurements with the iVue-100 optical coherence tomographer

**Background:** The reliability of an instrument is dependent on its accuracy and repeatability. The iVue-100 optical coherence tomographer (OCT) is a spectral domain device that has the advantage of being able to measure both the anterior and posterior segments of the eye. But its repeatability for macular thickness measurements has not been assessed and this was the primary purpose of this study.

**Aim:** To determine the repeatability of the iVue-100 OCT for macular thickness measurements.

**Settings:** This study was conducted at the University of KwaZulu-Natal.

**Methods:** Repeatability (intra-observer, inter-observer and inter-session) of the iVue-100 OCT was assessed on 50 participants selected using convenience sampling. Data were statistically analysed using Bland–Altman analysis, the intra-class correlation coefficient (ICC), coefficient of variation and the paired t-test.

**Results:** The mean age and standard deviation for the 50 participants were 23.88 ± 6.93 years (range 18–51 years). The ICCs for intra-observer repeatability ranged between 0.940 and 0.989, whereas the coefficients of variation ranged between 0.040 and 0.084. For both inter-observer and inter-session repeatability, the t-test showed no statistically significant differences (p > 0.05), and Bland–Altman analysis found the majority of the measurements to be within the 95% limits of agreement.

**Conclusion:** Irrespective of whether repeated measurements are being taken consecutively by the same examiner, by a different observer or at a different time, the iVue-100 OCT is highly reliable in terms of repeatability which is of paramount importance in any clinical or research setting.

**Introduction**

The word fovea refers to a small region and in terms of the retina is the area in the posterior pole measuring approximately 1.5 mm in diameter, situated approximately 4 mm temporal to the optic disc.\(^1,2\) The fovea allows the best visual resolution in the entire retina.\(^3\) Unlike the rest of the retina, the macula contains more than one layer of ganglion cells and an additional pigment called xanthophyll. Histologically, the macula is subdivided into the fovea, parafovea and perifovea from the centre to its periphery.\(^4\) The cone density decreases as one moves from the fovea to the perifovea while the rod density increases.\(^5\) The centre of the fovea, called the foveola, is completely devoid of both ganglion and inner nuclear cells but has the highest density of cones.\(^1,2,5\) The parafovea, located between the foveola and the perifovea, has fewer blood vessels but the highest density of ganglion cells. The perifovea, which is rich in blood vessels, is where the ganglion cell layer is the thickest.\(^6\) The histological differences of the different areas result in varying thickness in the sub-areas of the macula.

Macular thickness is least at the foveola, thickest within 3 mm of the centre and then decreases towards its periphery, with the temporal part being thinner than the nasal quadrant.\(^7\) Differences in macular thickness with gender and ethnicity have been reported.\(^8\) Ocular pathologies including retinal vein occlusion, diabetic retinopathy and age-related macular degeneration can result in changes in macular thickness.\(^9\) Traditional methods for measuring macular thickness include fluorescein angiography and slit lamp biomicroscopy with a retinal lens; however, these methods are qualitative.\(^10\) The advent of optical coherence tomography (OCT) has provided both quantitative and qualitative methods for measuring ocular structures with good precision.\(^9,11\)

Optical coherence tomography is an objective, non-invasive imaging technology,\(^12\) first introduced in 1991, and is similar to ultrasonography except that reflected light rays instead of sound waves...
are analysed.\textsuperscript{13,14,15} OCT is based on the principle of low-coherence interferometry\textsuperscript{16} whereby the reflected light rays are examined to create high-resolution cross-sectional images (tomograms). The two methods of detecting and examining the reflected light rays are referred to as time-domain OCT and Fourier-domain OCT.\textsuperscript{17} Traditional time-domain OCT devices have lower axial resolutions (8 μm – 10 μm) and slower scanning speeds (400 A-scans per second). The new-generation Fourier-domain OCT devices are capable of achieving scanning speeds of greater than 20 000 A-scans per second with 5 μm – 7 μm axial resolutions.\textsuperscript{17} Furthermore, Fourier-domain OCT devices are less affected by eye movements and allow for more stable tomograms\textsuperscript{17} with better visualisation of the retinal layers.\textsuperscript{19} Subsequently, OCT plays an important role in the screening, diagnosis and monitoring of retinal diseases.\textsuperscript{10,13,19}

Measurements of macular thickness with different OCT devices are not interchangeable,\textsuperscript{12} and because of their extensive use in monitoring pathologies, the reproducibility of measurements with each device is essential.\textsuperscript{10} Previous studies\textsuperscript{20,21,22} have found good repeatability with other OCT devices and concluded that they are effective in monitoring retinal thickness in the macular area particularly in diseased states. With technological devices, newer instruments are constantly being produced with claims of better optical resolution, higher speed and increased precision, which need to be verified. This study therefore set out to determine the repeatability of macular thickness measurements with the iVue-100 OCT which, to the best of the researcher’s knowledge, has not been studied previously.

Research methods and design

This was an observational cross-sectional study, and convenience sampling was used to recruit 50 healthy participants, of all races, gender and ages. The study adhered to the tenets of the Declaration of Helsinki, and all participants were required to give written informed consent. Participants recruited had normal corneal topography, visual acuity in each eye of at least 6/6 (aided or unaided) and no history of ocular injury and/or surgery.

The iVue-100 OCT device was used to scan and measure retinal thickness at the macula. This Fourier-domain OCT device has a scanning speed of 26 000 A-scans per second and axial resolution of 5 μm.\textsuperscript{23} The retinal map scan protocol in the iVue-100 OCT was used to determine macular thickness. This scan protocol consists of a raster pattern of 13 horizontal line scans of 512 A-scans and 7 horizontal line scans of 1024 A-scans within the central 1.5 mm vertical zone.\textsuperscript{23} The tomograms were automatically analysed (to segment the retinal layers) with the preprogrammed algorithm in the iVue-100 OCT device. This algorithm defines the retinal pigment epithelium and the inner limiting membrane as the outer and inner limits of the retina, respectively.\textsuperscript{23} Consequently, the retinal thickness is automatically calculated as the distance between these two limits.

When scanning, participants fixated an internal fixation target while the real-time image of the participant’s eye and corresponding retina was monitored on the laptop screen. Chin and forehead rests were used to stabilise the participant’s head. As per the manufacturer’s recommendations, repeat scans were taken if the scan quality index was less than 40 or labelled as ‘poor’ on the laptop screen display.\textsuperscript{23} The retinal scanning protocol was repeated three times per eye on the same participant to investigate intra-observer repeatability. A second observer repeated the retinal scanning protocol on the same participant at each visit to determine inter-observer repeatability. Observer one repeated the scanning protocol on the same participant on another day for inter-session repeatability. The approximate time between sessions ranged from 1 to 70 days. Each observer realigned the iVue-100 OCT device before capturing each scan. Both eyes of each participant were scanned with the measurements taken between 10h00 and 14h00. All macular thickness measurements were taken within the same session.

The retinal map scan protocol displays retinal thickness in a retinal map as defined by the Early Treatment Diabetic Retinopathy Study (ETDRS).\textsuperscript{24} This retinal map shows the average retinal thickness in the nine areas (Figure 1). Central foveal thickness refers to the average thickness in the central 1 mm ring of the ETDRS\textsuperscript{24} macular map. The inner and outer rings in the ETDRS\textsuperscript{24} macular map, of 3 mm and 5 mm diameter, correspond to the parafovea and perifovea, respectively (Figure 1). The parafovea and perifovea are divided into four quadrants (superior, inferior, nasal and temporal). The retinal map displays the average thicknesses of nine areas of the ETDRS\textsuperscript{24} macular map using a false-colour tomographic image display (Figure 1).

The Statistical Package for Social Sciences (SPSS version 24) was used to analyse the data. The intra-class correlation coefficient (ICC) and coefficient of variation (CoV) were used to report on the repeatability of the measurements. The ICC can range between zero and one and describes the correlation between repeated measurements. The higher the ICC, the better the repeatability. The CoV expresses the standard deviation in relation to the mean and can also range from one to zero, with a value closer to zero indicating better repeatability. In addition, Bland-Altman analysis and t-tests were used to report on the inter-observer and inter-session repeatability. The significance level was set at 95% with p-value of ≤ 0.05.

Ethical considerations

Data collection commenced after ethical clearance was obtained from the Biomedical Research and Ethics Committee (BE024/13).

Results

Demographics

The sample (N = 50) comprised of 64% (n = 32) women and 36% (n = 18) men with a mean age of 23.88 ± 6.93 years.
The racial distribution included 54% Indians, 36% Africans, 8% Europeans and 2% Asians. Fifty-two percent of the sample were emmetropic. The spherical equivalents for the right and left eyes, respectively, ranged from -8.38 to 1.63 D and -9.63 to 3.00 D. The right and left eye measurements were highly correlated ($r \geq 0.895$) in the nine areas with differences $\leq 2.30 \mu m$; therefore, in this article only the right eye measurements were analysed and reported upon.

### Intra-observer repeatability

Table 1 shows ICC together with the 95% confidence intervals (CIs) and the CoV for each observer at the different areas of the macula as measured.

The ICC for observer one ranged from 0.946 to 0.989 and that for observer two ranged from 0.940 to 0.988, indicating good repeatability. The CoV for observer one ranged from 0.040 to 0.081 and that of observer two from 0.042 to 0.084, again indicating excellent intra-observer repeatability.

### Inter-observer repeatability

Table 2 shows the mean differences and their standard deviations when comparing the measurements of each observer at the different macular areas for the 50 right eyes. Table 2 also shows the Bland-Altman limits of agreement. The $t$-test was used in the specific regions to determine the $r$- and $p$-values as an indication of significant differences between observers.
### Table 2: The mean differences ($\bar{X}$), and their standard deviations for macular thickness (µm) between observers, Bland-Altman upper and lower limits of agreement, t and p-values from paired t-tests and the intra-class correlation coefficients.

<table>
<thead>
<tr>
<th>Area of macula</th>
<th>$\bar{X}$ (µm)</th>
<th>s.d. (µm)</th>
<th>Upper LoA</th>
<th>Lower LoA</th>
<th>t</th>
<th>p</th>
<th>ICC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fovea</td>
<td>-0.227</td>
<td>5.522</td>
<td>10.596</td>
<td>-11.050</td>
<td>-0.290</td>
<td>0.773</td>
<td>0.981</td>
</tr>
<tr>
<td>Superior parafovea</td>
<td>-0.007</td>
<td>3.655</td>
<td>7.157</td>
<td>-7.171</td>
<td>-0.013</td>
<td>0.990</td>
<td>0.983</td>
</tr>
<tr>
<td>Inferior parafovea</td>
<td>0.260</td>
<td>5.191</td>
<td>10.434</td>
<td>9.914</td>
<td>0.354</td>
<td>0.725</td>
<td>0.961</td>
</tr>
<tr>
<td>Nasal parafovea</td>
<td>-1.120</td>
<td>4.543</td>
<td>7.784</td>
<td>-10.024</td>
<td>-1.743</td>
<td>0.088</td>
<td>0.977</td>
</tr>
<tr>
<td>Temporal parafovea</td>
<td>-0.940</td>
<td>3.515</td>
<td>5.949</td>
<td>-7.829</td>
<td>-1.891</td>
<td>0.065</td>
<td>0.982</td>
</tr>
<tr>
<td>Superior perifovea</td>
<td>-0.627</td>
<td>3.197</td>
<td>5.639</td>
<td>-6.893</td>
<td>-1.386</td>
<td>0.172</td>
<td>0.980</td>
</tr>
<tr>
<td>Inferior perifovea</td>
<td>-0.260</td>
<td>4.742</td>
<td>9.034</td>
<td>-9.554</td>
<td>-0.388</td>
<td>0.700</td>
<td>0.961</td>
</tr>
<tr>
<td>Nasal perifovea</td>
<td>-0.447</td>
<td>3.646</td>
<td>6.699</td>
<td>-7.593</td>
<td>-0.866</td>
<td>0.391</td>
<td>0.981</td>
</tr>
<tr>
<td>Temporal perifovea</td>
<td>0.300</td>
<td>3.767</td>
<td>7.683</td>
<td>-7.083</td>
<td>0.563</td>
<td>0.576</td>
<td>0.973</td>
</tr>
</tbody>
</table>

s.d., standard deviation; LoA, limits of agreement; ICC, intra-class correlation coefficients.

The solid line represents the $\bar{X}$ (-0.227 µm) and the dashed lines represent the upper and lower limits of agreement (see Table 2). An outlier near -30 µm has resulted in the width of limits of agreement being almost 22 µm.

**Figure 2**: A Bland–Altman plot comparing iVue-100 macular thicknesses of the 50 right eyes taken by observers one and two.

There was no statistically significant difference in means between the two observers for all nine macular areas ($p > 0.05$).

Overall, the measurements between the two observers were found to be within 1 µm of each other in all areas, with the exception of the nasal parafoveal area, where the mean difference was 1.120 µm. The ICC ranged between 0.961 and 0.983, indicating good repeatability.

Bland-Altman plots were used to graphically compare inter-observer macular thicknesses of all areas as taken by two observers. The Bland-Altman plot for the foveal macular thickness is illustrated in Figure 2.

With the exception of two measurements, all measurements were within the 95% limits of agreement.

**Inter-session repeatability**

Only the measurements of observer one were used to determine the inter-session repeatability. This observer had completed a higher number ($n = 33$) of inter-session repeat readings taken on two separate sessions. The interval between sessions 1 and 2 ranged between 1 and 70 days.

The mean difference of the macular readings ranged from 0.020 µm to 0.799 µm (Table 3) with the ICC ranging between 0.967 and 0.993. Paired t-tests were used to determine the t- and p-values as an indication of significant differences between sessions. For all regions, there were no statistically significant differences in the mean macular thickness between sessions (see Table 3).

Bland-Altman plots were used to graphically compare the macular thicknesses for all areas taken by observer one over two sessions, but only the Bland-Altman plot for the foveal macular thicknesses is illustrated here in Figure 3. With the exception of only two measurements, all measurements were within the 95% limits of agreement. The foveal mean difference (0.313 µm) was almost zero and less than 1 µm.

**Discussion**

In any clinical setting, it is imperative that the quantitative measurement of a variable is not only reliable but also reproducible. Accurate and consistent measurements of macular thickness are essential when screening for, diagnosing and/or monitoring macular diseases, as well as glaucomatous damage. For the iVue-100 OCT to be classified as producing valid clinical macular thickness measurements, its repeatability has to be high. To the best of our knowledge, this aspect of the iVue-100 OCT has not previously been measured.

The test–retest reliability of an instrument can be assessed in terms of intra-observer, inter-observer and inter-session repeatability. While intra-observer repeatability refers to consistency of measurements when taken by the same operator, inter-observer repeatability is assessed using multiple operators. Inter-session repeatability requires the comparison of measurements taken by a single operator but on two or more separate sessions. The reliability of an instrument is important for clinical and research applications. The ICC, CoV and Bland–Altman analysis have been used as indicators of instrument repeatability in previous studies involving macular thickness measurements.

The ICC was used to describe the correlation and relationship between repeated measurements for all three aspects, that is, intra-observer, inter-observer and inter-session repeatability. An ICC of 1 implies that the measurements are perfectly correlated. An ICC between 0.81 and 0.99 represents good
agreement between repeated measurements. The ICCs for intra-observer, inter-observer and inter-session repeatability were consistently above 0.90, ranging between 0.940 and 0.993, indicating good agreement for all three aspects and for all nine areas measured. Garcia-Martin et al. using a Cirrus Fourier-domain OCT also reported ICCs of higher than 0.9 for intra-session repeatability except for the temporal parafoveal (ICC of 0.881) and superior perifoveal (ICC of 0.832) areas.

The CoV that indicates the measurement variability in relation to the mean is also known as the relative standard deviation. It is obtained by dividing the standard deviation of the thickness measurements by its mean and is another helpful method of describing the reproducibility of repeated measurements. A CoV below 0.1 indicates high reproducibility whereas a CoV below 0.05 is considered as excellent. Intra-observer repeatability was assessed with the CoV in addition to the ICC. Excellent repeatability was obtained for all nine areas (range of 0.04–0.049) with the exception of the fovea, by both observers. A similar range was reported by Gürses-Özden et al. of 0.047–0.064 and Garcia-Martin et al. of 0.6–1.8 for macular thickness measurements using the time-domain OCT Model 3000 and Fourier-domain Cirrus OCT, respectively. In this study, the CoVs for the fovea were greater than 0.05 but less than 0.1 which is still regarded as good repeatability (CoV of observers 1 and 2 were 0.081 and 0.084, respectively). Other studies also reported more variability in their CoVs for foveal measurements. As the fovea (1.5 mm in diameter according to Kanski) comprises a small part of the macula, small shifts in alignment could bring the instrument onto the foveal margins which could lead to significant differences in thicknesses for this area and thus impact the repeatability.

Paired t-tests and Bland–Altman analysis were the third methods used in this study for both the inter-observer and inter-session repeatability. Overall, the measurements between the two observers were found to be within 1 µm of each other in all areas with the exception of the nasal parafoveal area where the difference was 1.120 µm. Similarly, the mean difference for all nine areas between sessions was also within 1 µm of each other. There were no statistically significant differences in measurements, for all nine areas between observers or between sessions. Garcia-Martin et al. also noted no significant differences between foveal thickness measurements taken in two sessions using the Cirrus OCT. In this study, the mean difference in foveal thicknesses for inter-session repeatability was only 0.313 µm. Garcia-Martin et al. noted a slightly higher difference of 1.9 µm in mean foveal thickness measurements taken over two sessions. Even though both the iVue-100 and the Cirrus OCT are Fourier-domain instruments, differences in the algorithms, scanning speeds and axial resolution may account for the variation in findings.

Bland–Altman analysis quantifies the agreement between two quantitative measurements using the mean difference and the construction of limits of agreement. Agreement between two measurements is concluded if 95% of the data points lie within two (1.96) standard deviations of the mean difference. In this study, good repeatability was obtained for both inter-observer and inter-session repeatability as 96% of data points were within the 95% limits of agreements for the foveal region. The limitation of this type of analysis is that it does not indicate if the limits are clinically meaningful or not. The limits of agreement for both observers and between sessions were 21.65 µm and 18.37 µm, respectively. Considering the average foveal thickness as measured with the iVue-100 in a myopic South African population was 240.50 µm, the limits of agreement in this study represent less than 9% of the thickness. If the outliers were removed or the sample size increased, then perhaps the LoA might reduce.

TABLE 3: The mean differences (\(\bar{X}\)) and their standard deviations for iVue-100 macular thickness (µm) for observer one taken over two sessions (between 1 and 70 days apart). Bland-Altman upper and lower limits of agreement, and t- and p-values from paired t-tests and the intra-class correlation coefficients.

<table>
<thead>
<tr>
<th>Area of macula</th>
<th>(X_{1}) (µm)</th>
<th>s.d. (µm)</th>
<th>Upper LoA</th>
<th>Lower LoA</th>
<th>t</th>
<th>p</th>
<th>ICC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fovea</td>
<td>0.313</td>
<td>4.687</td>
<td>9.500</td>
<td>-8.874</td>
<td>0.384</td>
<td>0.704</td>
<td>0.987</td>
</tr>
<tr>
<td>Superior parafoveal</td>
<td>-0.799</td>
<td>3.454</td>
<td>5.972</td>
<td>-7.568</td>
<td>-1.327</td>
<td>0.194</td>
<td>0.984</td>
</tr>
<tr>
<td>Inferior parafoveal</td>
<td>-0.374</td>
<td>4.113</td>
<td>7.687</td>
<td>-8.435</td>
<td>-0.522</td>
<td>0.605</td>
<td>0.974</td>
</tr>
<tr>
<td>Nasal parafoveal</td>
<td>-0.606</td>
<td>3.711</td>
<td>6.668</td>
<td>-7.880</td>
<td>-0.938</td>
<td>0.355</td>
<td>0.985</td>
</tr>
<tr>
<td>Temporal parafoveal</td>
<td>-0.465</td>
<td>2.342</td>
<td>4.125</td>
<td>-5.055</td>
<td>-1.140</td>
<td>0.263</td>
<td>0.992</td>
</tr>
<tr>
<td>Superior perifoveal</td>
<td>0.253</td>
<td>2.791</td>
<td>5.723</td>
<td>-5.217</td>
<td>0.520</td>
<td>0.607</td>
<td>0.986</td>
</tr>
<tr>
<td>Inferior perifoveal</td>
<td>-0.626</td>
<td>2.937</td>
<td>5.313</td>
<td>-6.383</td>
<td>-1.225</td>
<td>0.229</td>
<td>0.984</td>
</tr>
<tr>
<td>Nasal perifoveal</td>
<td>0.030</td>
<td>3.235</td>
<td>4.587</td>
<td>-4.527</td>
<td>0.075</td>
<td>0.941</td>
<td>0.993</td>
</tr>
<tr>
<td>Temporal perifoveal</td>
<td>0.020</td>
<td>4.143</td>
<td>8.140</td>
<td>-8.100</td>
<td>0.028</td>
<td>0.978</td>
<td>0.967</td>
</tr>
</tbody>
</table>

s.d., standard deviation; LoA, limits of agreement; ICC, intra-class correlation coefficients.

FIGURE 3: Bland-Altman plot comparing iVue-100 foveal macular thickness measurements of the 50 right eyes taken by observer one in two separate sessions separated by a minimum of 1 day and a maximum of 70 days.
Conclusion
The iVue-100 OCT demonstrates good intra-observer, inter-observer and inter-session repeatability for macular thickness measurements, which validates its use in a clinical setting and also for research. This study, however, did not determine the accuracy of the measurements. Future studies should evaluate the repeatability of this instrument in eyes with conditions such as macular pathology or glaucoma to more completely understand its reliability when measuring macular thicknesses.

Acknowledgements

Competing interests
The authors declare that they have no financial or personal relationships that may have inappropriately influenced them in writing this article.

Authors’ contributions
N.R. and R.H. conceptualised the study, collected and analysed the data, and wrote the manuscript.

References