Title: Comparison of reliability and repeatability of corneal curvature assessment with six keratometers

Author(s): Hamer, C., Buckhurst, H., Purslow, C., Shum, G., Habib, N. and Buckhurst, P.

Copyright, publisher and additional information:

This is the peer reviewed version of the following article: Hamer, C., Buckhurst, H., Purslow, C., Shum, G., Habib, N. and Buckhurst, P. (2016) Comparison of reliability and repeatability of corneal curvature assessment with six keratometers. Clinical and Experimental Optometry. Vol. 99, No. 6: 583-589

which has been published in final form at http://doi.org/10.1111/cxo.12329. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Self-Archiving.

DOI: http://doi.org/10.1111/cxo.12329
Comparison of reliability and repeatability of corneal curvature assessment with six keratometers

Running title: Comparison of Six Keratometers

Authors: Catriona A Hamer\textsuperscript{a} BSc (Hons) MCOptom, Hetal Buckhurst\textsuperscript{a} PhD MCOptom, Christine Purlsow\textsuperscript{b} PhD MCOptom, Gary L Shum\textsuperscript{a} PhD, Nabil E Habib\textsuperscript{c} FRCS FROphth, Phillip J Buckhurst\textsuperscript{a} PhD MCOptom

Author Institutions:

\textsuperscript{a}School of Health Professions, Faculty of Heath and Human Sciences, Plymouth University, Plymouth, UK

\textsuperscript{b}School of Optometry & Vision Sciences, Cardiff University, UK

\textsuperscript{c}Royal Eye Infirmary, Derriford Road, Plymouth, UK

Corresponding author:

Dr Phillip Buckhurst
School of Health Professions, Peninsula Allied Health Centre, Plymouth University, Plymouth, UK, PL6 8BH
Tel: +44(0) 1752 588 884
Email: phillip.buckhurst@plymouth.ac.uk

Financial disclosure:
Consultant for Bausch and Lomb

[Type text]
Subtitle:
The agreement and reproducibility of 6 keratometry instruments is investigated. When considering mean spherical error alone, IOLMaster, Pentacam, OPD-scan and Medmont may be considered interchangeable however astigmatism shows greater variability between instruments, sessions and observers.
(Abstract:)

**Background/ Aims**

Keratometry methodology varies between instruments and the differences may potentially have a clinical impact. We investigated the agreement and reproducibility of six keratometry instruments.

**Methods**

Keratometry was performed on 100 subjects at two separate sessions with IOLMaster 500, Pentacam, OPD scanner, Medmont E300, Javal Schiøtz and TMS-5. A second observer assessed 30 subjects to determine inter-observer variability. A single individual was assessed on 10 separate sessions to determine intra-observer variability. Data were analysed using coefficient of variation (CV) and intra-class correlation coefficient (ICC) for intra-observer variation. Inter-observer concordance was evaluated by ICC. Bland-Altman plots, Pearson’s correlation coefficient and repeated measures ANOVA were used to assess agreement of data produced by the instruments.

**Results**

OPD scanner and Javal Schiøtz mean spherical equivalent (MSE) results were systematically different ($p < 0.001$) to other instruments (flatter and steeper, respectively). $J_0/J_{45}$ were similar for all instruments ($p < 0.05$). Bland-Altman comparison plots indicated that Pentacam and IOLMaster demonstrated greatest level of agreement (ICC results MSE = 0.992, $J_0$ = 0.934 and $J_{45}$ = 0.890)
Agreement (ICC) between observers for MSE ranged from 0.955-0.995 for all instruments; lower levels of agreement were found for J₀/J₄₅ (0.289-0.901). IOLMaster showed greatest correlation and Medmont the lowest. All instruments showed high intra-observer repeatability of MSE (CoV 0.1-0.3 %). The J₀J₄₅ readings showed greater variability (COV range 8.8 - 57.6 %).

**Conclusion**

When considering MSE alone IOLMaster, Pentacam, OPD scan and Medmont may be considered interchangeable. However, assessment of astigmatism shows greater variability between instruments, sessions and observers.

Keywords: Keratometry, corneal topography, corneal curvature, astigmatism
INTRODUCTION

The cornea is responsible for 2/3 of the total optical power of the eye and hence the precise assessment of corneal curvature prior to cataract surgery is vital for achieving optimal refractive outcomes. Until recently the importance of accurately determining both corneal astigmatic power and axis has been under-utilized, since spherical intraocular lens (IOL) power determination is based on average corneal curvature. In recent years toric IOLs have become an increasingly popular choice for the correction of astigmatism.\(^1\) The effectiveness of a toric IOL is dependent on its orientation and power in relation to the corneal principal meridians.\(^2\) Consequently the importance of reliably identifying and assessing the principal corneal meridians of curvature (power) has been highlighted.

Numerous instruments are commercially available for the assessment of corneal curvature and the outcomes of these instruments are widely considered to be interchangeable.\(^3\)\(^-\)\(^5\) However, given that the optical principles behind these instruments differ, it is likely that inherent differences between devices exist when assessing corneal power. Furthermore, in much of the published literature examining the validity and repeatability of these instruments the emphasis has been on the mean spherical curvature alone ignoring the accuracy of the astigmatic orientation and magnitude been analysed in detail. A recent trend given the popularity of toric IOLs is to examine corneal curvature through vector analysis \(^6\)\(^-\)\(^11\) as this provides a more detailed and relevant assessment of corneal power.\(^12\)\(^,\)\(^13\)
The primary goal of this study was to assess the variability, reliability and agreement of corneal curvature measures determined with a range of commercially available devices.

**Methods and Materials**

One hundred adult subjects (32 Males, 68 Females) were recruited from the Plymouth University staff and student population. All procedures followed the Declaration of Helsinki and the protocol was reviewed and approved by the Plymouth University Ethics committee. The mean age was 36.0 ± 11.4yrs, ranging 19 - 57 years old. The inclusion criteria required each subject to be a consenting adult aged 18 and over, with healthy corneas. The exclusion criteria included previous refractive or other corneal surgery, RGP contact lens wear, corneal dystrophies or other abnormal corneal pathology. Soft contact lenses wearers were asked not to wear their contact lenses on the day of assessment with at least 12 hours since last wear.

Corneal curvature was recorded in all 100 subjects with six instruments, each calibrated at the beginning and at set intervals of the study, in a randomised order:

**Javal-Schiøtz**

Keratometry utilizes the principals of reflection; the corneal surface and tear film act as a convex mirror, which reflect the image of an object at a given distance. The curvature of the cornea is then determined through analysis of the resultant image. Keratometry assumes a spherical corneal shape and is highly dependent on a stable tear film. The Javal-Schiøtz is a two–position, fixed
doubling, manual keratometer and calculates the corneal curvature over a 3.4 mm diameter area.\textsuperscript{14}

IOLMaster 500 (Carl Zeiss Meditec Inc., Jena, Germany)
The IOLMaster utilizes automated keratometry for the assessment of corneal curvature. It projects 6 spots in a hexagonal pattern of light onto the corneal/tear film at a diameter less than 2.3 mm. The separation of the opposite pairs of lights is measured objectively by the instrument’s internal software. In the case of an astigmatic cornea, the curvature is calculated from three, fixed position meridians.\textsuperscript{11}

Pentacam (Oculus Optikgeräte GmbH, Wetzlar, Germany)
The \textit{Pentacam HD} is a non-contact anterior segment imaging device that is based on the principles of rotating Scheimpflug photography. The instrument uses a monochromatic slit light source (i.e. a blue LED at 475 nm) and a Scheimpflug camera, which together rotate around the optical axis of the eye.\textsuperscript{15}
The Simulated K readings (based on anterior corneal curvature alone) can be obtained over a small central area (3 mm) that allows comparison with other instruments.

OPD scanner (Nidek Co., Ltd, Gamagori, Japan)
The OPD scanner III assesses corneal curvature using computerised placido disc topography, again utilising principals of reflection. A placido disc is projected onto the cornea/tear film, and the computer then analyses thousands of points reflected from the whole cornea. It simulates Ks over a 3 mm area.\textsuperscript{16}

Medmont E300 (Medmont PTY Ltd., Camberwell, Victoria, Australia)
The Medmont is a computerised placido disc cone videokeratometer. It has 32 placido rings and measures 9,600 data points per scan. The simulated K-readings are the steep and flat radius of curvature found over a 3 mm area.\(^5\)

The TMS-5 (topographical modelling system, \textit{TOMEY Corp., Nagoya, Japan})

The TMS-5 incorporates both a 31-ring placido disc topographer and Scheimpflug tomographer. The results from the Scheimpflug measurement and topographical measurement are combined to produce an adjusted measurement.\(^6,17\)

Each subject was assessed on two separate sessions by a single trained, observer; a subgroup of 30 subjects were assessed again with each instrument by a second trained observer within the second session to determine the inter-observer variability; the second observer being blind to the results from the first observer. A single randomly selected subject was assessed on 10 separate measurement sessions (separated by a minimum of 24 hours) by a single observer to determine the intra-observer variability for each instrument.

Statistical Analysis

The size of the subject group was determined with an alpha level of 0.05 and a power of 80% confidence. Multiple sample size test calculations were carried out with the G*Power 3 (Heinrich Heine Universität, Düsseldorf, Germany) programme to determine the size with the assumption of a moderate effect size advised by Cohen’s table comparison of paired means (effect size 0.50), correlation (effect size 0.30) and ANOVA analysis (effect size 0.25). A minimum sample size of 84 was required to satisfy power requirements across these
analyses. Therefore, 100 volunteers were recruited to allow for drop outs or exclusion of some subjects throughout the study. For the selection of the subgroup a minimum of 22 subjects were needed to compare the two observers; 30 subjects were recruited to allow for any loss throughout the study. Bland-Altman plots were created to assess the agreement between the machines showing the mean and 95% limits of agreement (mean ±1.96 DS).

The data was analysed using SPSS software (Version 20, SPSSInc, IBM, Chicago, Illinois, USA), and tested for normality using a Kolmogorov-Smirnov test prior to statistical analysis.

All keratometry results were converted to rectangular Fourier form of mean spherical equivalent (MSE) and \(J_0/J_{45}\) representing the cylindrical power and axis as a combined vector for analysis. The MSE was calculated by adding half the cylindrical power to the spherical power: \(\text{MSE} = \text{Sph} + \frac{1}{2} \text{Cyl}\). \(J_0\) and \(J_{45}\) were converted into vectors using the following formulae:

\[
J_0 = J \cos (2\alpha) \\
J_{45} = J \sin (2\alpha) 
\]

Inter-observer repeatability was assessed via the Intra-Class Correlation Coefficient (ICC) on a subgroup of 30 subjects who had been assessed by two separate examiners. Intra-observer repeatability was determined by examining the Coefficient of Variance (CoV) for the single subject who had been assessed 10 times on each device and by examining the Intra-Class Correlation (ICC) on
100 subjects where the measurements were performed twice by a single examiner. Bland-Altman plots were created with SigmaPlot (SYSTAT software Inc, San Jose, California, USA) and used to determine the agreement and therefore potential inter-changeability of the instruments.

 Pearson’s correlations were used to determine the correlation of results between instruments. The difference between means was assessed using repeated measures ANOVA followed by a Bonferroni post-hoc test on the results shown to be significant.

**RESULTS**

The Kolomogorov-Smirnov test found the data to be normally distributed (p>0.005).

**Inter-observer repeatability**

The inter-observer repeatability for MSE (Table 1) was greater than 0.95 for all instruments. The inter-observer repeatability for J₀/J₄₅ (Table 1) showed greater variability than that for MSE particularly in respect to the Medmont and Javal Schiøtz (Table 1). The Pentacam and IOLMaster demonstrated the greatest inter-observer repeatability.

The ICC between visits for all 100 subjects (Table 1) shows similar pattern of results as Table 1 and 2. The Pentacam showed the highest correlation whilst the TMS-5 showed the lowest.
Table 1: ICC between two observers for the second visit (n=30) and between visits for all 6 Instruments (n=100)

<table>
<thead>
<tr>
<th></th>
<th>ICC (between two observers) (n=30)</th>
<th>ICC (between two visits) (n=100)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MSE</td>
<td>MSE</td>
</tr>
<tr>
<td></td>
<td>IOLMaster</td>
<td>Pentacam</td>
</tr>
<tr>
<td>MSE</td>
<td>0.994</td>
<td>0.996</td>
</tr>
<tr>
<td>J0</td>
<td>0.901</td>
<td>0.933</td>
</tr>
<tr>
<td>J45</td>
<td>0.895</td>
<td>0.872</td>
</tr>
<tr>
<td>MSE</td>
<td>0.991</td>
<td>0.981</td>
</tr>
<tr>
<td>J0</td>
<td>0.829</td>
<td>0.911</td>
</tr>
<tr>
<td>J45</td>
<td>0.903</td>
<td>0.870</td>
</tr>
</tbody>
</table>

Intra-observer repeatability

The intra-observer repeatability (CoV) for MSE (Table 2) was less than 0.4 for all instruments with the IOLMaster showing least variation between readings by the same observer. In contrast, the intra-observer repeatability of J0/J45 (Table
2) showed much greater variability, particularly for the TMS-5 and Javal Schiøtz. The Pentacam and IOLMaster performed the best for J₀/J₄₅.

**Table 2: CoV (%) for all 6 Instruments (n=1)**

<table>
<thead>
<tr>
<th></th>
<th>IOLMaster</th>
<th>Pentacam</th>
<th>OPD</th>
<th>Medmont</th>
<th>Javal Schiøtz</th>
<th>TMS-5</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSE</td>
<td>0.1</td>
<td>0.3</td>
<td>0.3</td>
<td>0.2</td>
<td>0.3</td>
<td>0.2</td>
</tr>
<tr>
<td>J₀</td>
<td>11.0</td>
<td>8.8</td>
<td>17.6</td>
<td>18.0</td>
<td>23.8</td>
<td>31.2</td>
</tr>
<tr>
<td>J₄₅</td>
<td>6.9</td>
<td>8.7</td>
<td>24.1</td>
<td>32.6</td>
<td>57.6</td>
<td>49.3</td>
</tr>
</tbody>
</table>

**Mean vs. difference plots**

Bland-Altman comparison plots (Figures 1 and 2) indicated that the Pentacam and IOLMaster showed the greatest level of agreement for both MSE and J₀/J₄₅.

When assessing MSE the TMS-5 and Javal Schiøtz demonstrated the widest limits of agreement but when examining J₀/J₄₅ the OPD scanner showed poorest agreement.
Figure 1: Bland and Altman comparison of $J_0$ and $J_{45}$ for all pairs
Figure 2: Bland and Altman comparison of MSE for all pairs

[Type text]
Correlation

A strong positive correlation was found in the comparison of the MSE results for all combinations of pairs across the 6 instruments ($r = 0.888-0.922$, $p<0.001$). The correlation was strongest between Pentacam and IOLMaster ($r = 0.992$, $p<0.01$) and weakest between the TMS-5 and Javal Schiotz ($r = 0.888$, $p < 0.01$). Comparison of the results for the vertical and horizontal corneal astigmatism showed weaker correlation and more variability depending upon which pairings were assessed. The IOLMaster $J_0$ and $J_{45}$ values show the strongest correlation when compared to the Pentacam ($J_0$: $r = 0.934$, $p<0.001$; $J_{45}$: $r = 0.890$, $p<0.01$). There was a much weaker correlation between the IOLMaster and the OPD ($J_0$: $r = 0.720$, $p<0.001$; $J_{45}$: $r = 0.738$, $p<0.001$), Medmont ($J_0$: $r = 0.642$, $p<0.001$; $J_{45}$: $r = 0.835$, $p<0.001$), Javal Schiotz ($J_0$: $r = 0.0747$, $p<0.01$; $J_{45}$: $r = 0.531$, $p<0.001$) and TMS-5 ($J_0$: $r = 0.648$, $p<0.001$; $J_{45}$: $r = 0.5740$, $p<0.001$).

Comparison of means

Repeated measures ANOVA demonstrated a significant difference in MSE ($F = 84.977$, $p<0.001$). Post-hoc analysis showed that the OPD scan results were significantly different from the other instruments, finding a lower MSE (flatter cornea) on average ($p<0.01$ in all cases). In comparison the Javal Shiøtz also showed significantly different MSE ($p<0.001$) showing a higher average MSE (steeper cornea) in comparison with the other devices. In addition the MSE measurement was significantly steeper with the Medmont when compared to the Pentacam ($p = 0.01$) (Table 3).
There were no significant differences between any of the average measures of the astigmatic vector components ($J_0$: $F = 1.047$, $p = 0.372$; $J_{45}$: $F = 1.210$, $p = 0.307$) (Table 3).

**DISCUSSION**

Precise corneal astigmatism assessment is essential when choosing the power of an IOL to be implanted in cataract surgery. When using toric IOLs, a higher degree of accuracy is required to ensure that not only the power but also the orientation of the lens is positioned accurately to provide the optimum correction. The purpose of this study was to determine the repeatability and
validity of 6 different instruments designed to measure corneal curvature in pre-
surgical assessment.

In conventional cataract surgery, with non-toric IOLs, only the accuracy of MSE
is the important outcome when assessing corneal curvature. It was unsurprising
that all of the devices in the study demonstrated high MSE inter- and intra-
repeatability.

However, discrepancies were found between the MSE results when comparing
the instruments. This variation may be due to differences in the optical and
mathematical methods used to calculate corneal power. It was found that the
manual keratometer provided a steeper MSE than the other instruments; the
instruments that calculated Sim-K from placido disc topography provided the
flattest measurements; and the instruments that determined corneal curvature
through automated keratometry or Scheimpflug imaging provided results flatter
than manual keratometry but steeper than placido disc topography.

Previous studies evaluating the results of the Javal Schiøtz with the IOL Master
found that both provided similar results for MSE. 11 However, in the current
study the manual keratometer was found to measure steeper than all other
devices. The discrepancy found in the current study may be due to several
factors. The manual keratometers results are formed from an estimation of
corneal curvature based on the central 3.2mm zone as opposed to the central
2.3mm zone of the IOL Master. Unlike corneal topography and tomography the
manual keratometer assumes that the cornea is spherical in shape and cannot
determine an aspheric profile. Furthermore the manual keratometer has an inherent dependency on the examiner to accurately determine the end points.

In the current study the IOLMaster provided a steeper corneal curvature than the placido-disc topographers. These finding are in agreement with previous reports where the discrepancies in measures have been attributed to the small area, which it uses to simulate the K readings. 20,21

Previous studies examining the validity of the Pentacam reported that it produced systematically flatter corneal curvature readings than other instruments.9,20 However, these studies used the net corneal power measurement rather than anterior corneal curvature of the Pentacam and hence the results are not comparable. Reuland and associates used only the anterior corneal curvature for assessment and found that the IOL Master and Pentacam showed comparable results.22 In comparison Savani and colleagues found that the Pentacam measured a flatter anterior corneal curvature however, the authors used an older version of the Pentacam with a 25 scan setting.23

The analysis of corneal astigmatism measurement separate to the MSE highlights the difficulty of accurately determining astigmatic power and orientation. The assessment of MSE is not dependent on the orientation of the power meridians and is more robust to erroneous readings effecting one meridian. In this study the J_0 and J_45 vectors described by Thibos.24 are used for statistical analysis; this vector analysis allows the comparison of both orientation and power. Not all studies assessing keratometry assess this component separate to the MSE. 4,5,20,21,25
In comparison with the topographers and manual keratometer, the Pentacam and IOL Master demonstrated high repeatability between observers, visits and within observer repeatability for \( J_0/J_{45} \) (tables 2 and 3) and still show a very good agreement with Bland-Altman plots (figure 2). This is similar to previous studies whereby the Pentacam and IOL Master demonstrate high intra-observer, inter-observer and between session repeatability for \( J_0/J_{45} \).\(^3,^{11,26}\) The repeatability was weaker for those instruments based upon a topographic optical technique; this agrees with the findings of Wang and colleagues\(^5\) who found that there was a much larger spread in results and poorer repeatability with such instruments.

Unlike Scheimpflug imaging, it is likely that the tear film has a significant influence on the repeatability of topographic keratometers. It can be proposed that the tear film has a larger influence on the assessment of astigmatism than it does on average corneal power due to the influence of localised changes to the tear film. As astigmatism is orientation specific, a localised disturbance to the tear film can influence readings along a specific meridian and hence distort the measurement of astigmatism. When assessing the concordance of devices, an interesting observation is that those instruments based around placido disc cornea topography have produced a wider spread of data and more outliers in relation to the scheimpflug and automated keratometry techniques. This could provide further support to the influence of an unstable tear film creating disparate results and clear outliers in the data. The use of ocular lubricants prior to measurements may provide a more stable reading.
The limits of agreement shown by the Bland Altman plots demonstrate the disparity of agreement between various pairings of the instruments. If a disagreement of 0.50D is considered to be of clinical significance then the present results would suggest that the use of certain instruments in combination could lead to a significant under or overestimation of corneal power analysis. The MSE comparison of the topographers especially the TMS-5 and other instruments showed limits of agreement over 0.50D as demonstrated in Figure 1: F, I, L, N and O. Furthermore the comparison of the $J_0/J_{45}$ vector components in Figure 2 displays a similar increase in discrepancy for the TMS-5.

The IOLMaster 500 provided repeatable readings of $J_0/J_{45}$, which were similar to those of the Pentacam. The smaller measurement zone of 2.3 mm is likely to be an important factor as there is less chance of the measures being influenced by more peripheral tear film changes; furthermore the integrated software of the IOLMaster 500 has numerous image quality checks that may further improve the reliability of the measurement.

Previous work has also shown the IOLMaster astigmatism assessment to be interchangeable with the Javal Schiøtz $^{11}$ and the Pentacam with the Medmont.$^3$ In contrast to this; our study has shown much poorer agreement and repeatability when considering corneal astigmatism assessment with any of the other four devices. The intra-observer, inter-observer and inter-session...
repeatability are all much lower for the Medmont, OPD scanner, Javal Schiøtz and TMS-5 when assessing astigmatism.

This study had some limitations in design. The subjects who routinely wore soft contact lenses were advised to remove the lenses a minimum of 12 hours before the assessment. On review of the literature it appears that in some cases the corneal shape can be affected by soft contact lens wear for up to 2 weeks and thus a longer time period between cessation of contact lens wear and assessment would be advised to increase the accuracy of the readings. In the present study only 8 subjects were soft contact lens wearers (2 were infrequent wearers), limiting the affect of possible corneal changes on the study results. Although the study was performed on healthy phakic subjects (18-60), the present results provided an indication of the repeatability validity and concordance of results predicted for an older subject group such those having cataract surgery.

In conclusion, compared to MSE the variability between instruments is much greater when assessing corneal astigmatism. The Pentacam and IOLMaster appear to be the best choice for use with toric cataract surgery assessments however this requires further investigation in the post-operative environment. Specifically, future work needs to investigate the use of these two instruments in assessing the influence of corneal astigmatism on the ocular refraction in pseudophakic population.
REFERENCES


ACKNOWLEDGEMENTS

Thanks to Abby Pearce and Adam Carter in their technical assistance in this study.

Funding Sources: This study has been funded by Plymouth University as part of the first author’s PhD.

Potential conflict of interest: There were no conflicts of interests associated with any of the authors or their study.