Clinical evaluation of keratometry and computerised videokeratography: intraobserver and interobserver variability on normal and astigmatic corneas

Constantinos H Karabatsas, Stuart D Cook, John Papaefthymiou, Phil Turner, John M Sparrow

Abstract
Aims—To evaluate intra- and interobserver variability in measurements on normal and astigmatic corneas with keratometry and computerised videokeratography.

Methods—Keratometric readings with the 10 SL/O Zeiss keratometer and topographic maps with the TMS-1 were obtained by two independent examiners on 32 normal and 33 postkeratoplasty corneas. Inter- and intraobserver coefficients of variability (COR) for measurements of steep and flat meridian power and location, in addition to the magnitude of astigmatism, were assessed.

Results—Compared with TMS-1, the 10 SL/O keratometer showed a superior repeatability in measuring normal corneas (intraobserver COR for keratometry and TMS-1 respectively: 0.22 and 0.30 D for steep meridian power; 0.18 and 0.44 D for flat meridian power; 0.26 and 0.40 D for astigmatism; 5° and 26° for steep meridian location; 5° and 13° for flat meridian location). Astigmatism intraobserver COR (0.20 D and 0.26 D for the two observers) and interobserver COR (0.28 D) of the keratometer for normal corneas was very good and not affected by observers’ experience. Repeatability of the TMS-1 on normal corneas was found to be: (a) observer related, and (b) astigmatism related. A novice observer showed a much greater COR (1.62 D for astigmatism, 30° for flat meridian location) compared with the experienced examiner (0.40 D for astigmatism, 13° for flat meridian location). Higher deviation scores were observed for corneas with higher astigmatism. For the postkeratoplasty corneas, the keratometer achieved superior reproducibility (astigmatism intraobserver COR 1.12 D for keratometry, 4.06 D for TMS-1; steep meridian location interobserver COR 10° for keratometry, 34° for TMS-1).

Conclusion—Keratometric readings are more reproducible than topographic data both for normal and postkeratoplasty corneas. The two instruments should not be used interchangeably especially on highly astigmatic corneas. For the TMS-1, users with the same level of experience should be employed in clinical or experimental studies.

Measurement is a scientific method used in every aspect of clinical or experimental medicine. The terms precision, reproducibility, repeatability, or variability of an instrument refer to the agreement between repeated observations with the given instrument. On the other hand, accuracy of a method refers to how close the measured values are to the real value. Computer assisted videokeratoscopes (CAVK) are instruments that measure the corneal slope, calculate the corneal power, and reproduce the corneal profile with a certain level of accuracy and precision. It is important, and clinically useful, to know the reproducibility of these relatively new instruments, especially in comparison with the previously available “gold standard” ones such as the keratometer.

Previous studies have shown an acceptable level of accuracy and reproducibility for different commercially available models of CAVK, mainly on calibrated spherical test surfaces and normal corneas. However, in practical clinical terms we often use both keratometry and CAVK when measuring abnormal corneas such as those seen after penetrating keratoplasty (PKP). It would, therefore, be useful to know and compare the performance of these instruments not only in normal non-astigmatic corneas, but also in postoperative, highly astigmatic corneas.

A way of assessing the precision of instruments is either by examining the measurement agreement, or by looking at the intra- and interobserver variability of measurements. We have previously reported on the results of measurement agreement between keratometry and CAVK on both normal and post-PKP corneas. The aim of the present study was to assess the intraobserver and interobserver variability for both the keratometer and CAVK in normal as well as postkeratoplasty corneas. Evaluation of the variability was conducted in terms of dioptric power and axis of the two principal corneal meridians, as well as magnitude of corneal astigmatism.

Materials and methods

INSTRUMENTS
The keratometer used in this study was the 10 SL/O model (Carl Zeiss Ltd) which is an ophthalmometer after Helmholtz and follows the variable doubling principle of operation. The measured range of the instrument for radii of curvatures is from 4.00 to 11.2 mm, with a scale interval of 0.01 mm. The measured
corneal diameter is from 1.5 to 3.5 mm depending on the radius of the examined cornea. The scale accuracy throughout the entire measuring range is plus or minus $2 \times 10^{-2}$ mm. This is within the recommended tolerance for keratometers. The Zeiss ophthalmometer is a “two position” instrument and can thus measure principal corneal meridians that are not perpendicular. After identifying the radius of curvature of one principal meridian, the mires are rotated to locate and measure the other meridian which will be at an angle other than 90° if the astigmatism is irregular.

For the purpose of this study, the TMS-1 (Computed Anatomy, New York, NY, software version 1.61) model of CAVK with a 25 ring light cone was used. The system uses a short working distance, approximately 40 mm from the centre of the cone to the surface of the eye. The patient fixes a light centred in the cone and focusing of the instrument is done by manoeuvring the Placido target with a joystick. The instrument provides a non-spherocylindrical simulated keratometric (simK) value, which takes into consideration the power and axis of the flattest meridian regardless of the angle between the steepest and flattest meridian. In this study the non-spherocylindrical simK value was used in all calculations.

PATIENTS

For the normal population study, both corneas from 17 normal subjects (nine females) aged 18 to 64 (mean 36) years were studied. Corneas were considered normal and included in the study, only if there was: (1) no history of ocular surgery, (2) no slit-lamp microscopy evidence of trauma or corneal disease, (3) best corrected visual acuity (BCVA) of 6/60 or better to allow adequate fixation, (4) regular keratometric readings, (5) keratometric astigmatism of $<1.50$ D. Two of the 34 eyes had to be excluded from the study (one due to a central corneal opacity of unknown aetiology which produced distorted keratometric mires, and the second because of excess corneal astigmatism). This left 32 normal corneas for analysis.

For the astigmatic group, 33 postkeratoplasty corneas from 27 patients (15 females) aged 19 to 92 years (mean 53.2 years) were examined. Preoperative diagnosis included 12 eyes with keratoconus, nine eyes with Fuchs’ endothelial dystrophy, six eyes with herpetic keratitis, two eyes with aphakic bullous keratopathy, three eyes with various corneal dystrophies, and one eye was operated for keratoglobus. At the time of examination, all cases were within 12 months from the PKP (mean 7.1 months, range 1–12 months). At the time of examination, patients had a variable number and location of sutures present, or all sutures had already been removed. As with the normal population group, BCVA of 6/60 or better to allow adequate fixation with keratometry and CAVK was one of the entry criteria.

EXAMINERS AND EXAMINATION CONDITIONS

A total of three investigators were involved in the measurements, but as a pair for each of the two studies. Observers 1 and 3 were experienced in the use of both instruments. Observer 2 was a novice with the use of TMS-1, as well as with this particular keratometer model. All measurements for a given instrument were made independently by two operators in a masked fashion. For the normal corneas study observers 1 and 2 were employed; for the post-PKP group observers 1 and 3 were used.

All corneal topographic examinations were conducted in the same site. Keratometric readings were obtained in a similar room and the same keratometer was used in all cases.

Both instruments were calibrated before the start of the study.

Each observer obtained three measurements from each cornea with either the keratometer or the TMS-1. No artificial tears were used in any case. The three measurements were obtained after patients were instructed to blink and refixate between measurements, but without moving their head from the head rest. The sequence of the measurements with the two instruments was not randomised, but was not always done in the same order.

Before obtaining each measurement by the keratometer, the eyepieces of the instrument were set by each investigator to correct for their refractive error if they were not wearing contact lenses or spectacles. None of the observers was wearing astigmatic correction. Millimetre radius readings were taken from the radius of curvature scale of the 10 SL/O keratometer (our instrument did not have the facility of direct dioptre readings). A refractive index of 1.3375 was assumed for conversion from sphere radius to dioptres instead of the index
Table 1  Repeatability of the 10 SL/O Zeiss keratometer and TMS-1 videokeratoscope on post-PKP corneas

<table>
<thead>
<tr>
<th>Observer</th>
<th>Variance Mean (SD)</th>
<th>Intraobserver COR Mean (SD)</th>
<th>Variance Mean (SD)</th>
<th>Intraobserver COR Mean (SD)</th>
<th>Variance Mean (SD)</th>
<th>Intraobserver COR Mean (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Steep axis power (D):</td>
<td>0.02 (0.11)</td>
<td>0.22 D</td>
<td>0 (0.09)</td>
<td>0.18 D</td>
<td>0* (0.12)</td>
</tr>
<tr>
<td></td>
<td>Flat axis power (D):</td>
<td>0.07 (0.15)</td>
<td>0.30 D</td>
<td>-0.05 (0.54)</td>
<td>1.08 D</td>
<td>0.009 (0.46)</td>
</tr>
<tr>
<td></td>
<td>Astigmatism (D):</td>
<td>0.01 (0.09)</td>
<td>0.18 D</td>
<td>0* (0.08)</td>
<td>0.16 D</td>
<td>0.009 (0.10)</td>
</tr>
<tr>
<td></td>
<td>Flat axis power (D):</td>
<td>0.08 (0.22)</td>
<td>0.44 D</td>
<td>0.26 (0.98)</td>
<td>1.96 D</td>
<td>-0.17 (0.91)</td>
</tr>
<tr>
<td></td>
<td>Astigmatism (D):</td>
<td>0.01 (0.13)</td>
<td>0.26 D</td>
<td>0.002 (0.10)</td>
<td>0.20 D</td>
<td>0.00 (0.14)</td>
</tr>
<tr>
<td></td>
<td>Flat axis power (D):</td>
<td>0.00 (0.20)</td>
<td>0.40 D</td>
<td>-0.20 (0.81)</td>
<td>1.62 D</td>
<td>-0.10 (0.63)</td>
</tr>
<tr>
<td></td>
<td>Astigmatism (D):</td>
<td>3° (3°)</td>
<td>6°</td>
<td>3° (3°)</td>
<td>6°</td>
<td>4° (4°)</td>
</tr>
<tr>
<td></td>
<td>Flat axis power (D):</td>
<td>8° (13°)</td>
<td>26°</td>
<td>11° (11°)</td>
<td>22°</td>
<td>15° (20°)</td>
</tr>
<tr>
<td></td>
<td>Astigmatism (D):</td>
<td>3° (3°)</td>
<td>6°</td>
<td>3° (4°)</td>
<td>8°</td>
<td>5° (4°)</td>
</tr>
<tr>
<td></td>
<td>Flat axis power (D):</td>
<td>12° (7°)</td>
<td>13°</td>
<td>7° (15°)</td>
<td>30°</td>
<td>14° (21°)</td>
</tr>
</tbody>
</table>

Coefficient of repeatability (COR) calculated as 2 × SD.

Table 2  Repeatability of the 10 SL/O Zeiss keratometer and TMS-1 videokeratoscope on post-PKP corneas

<table>
<thead>
<tr>
<th>Observer</th>
<th>Variance Mean (SD)</th>
<th>Intraobserver COR Mean (SD)</th>
<th>Variance Mean (SD)</th>
<th>Intraobserver COR Mean (SD)</th>
<th>Variance Mean (SD)</th>
<th>Intraobserver COR Mean (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Steep axis power (D):</td>
<td>-0.05 (0.44)</td>
<td>0.88 D</td>
<td>-0.13 (0.47)</td>
<td>0.94 D</td>
<td>-0.06 (0.52)</td>
</tr>
<tr>
<td></td>
<td>Flat axis power (D):</td>
<td>0.09 (0.71)</td>
<td>1.42 D</td>
<td>0.015 (0.91)</td>
<td>1.82 D</td>
<td>0.05 (0.94)</td>
</tr>
<tr>
<td></td>
<td>Astigmatism (D):</td>
<td>-0.11 (0.66)</td>
<td>1.32 D</td>
<td>0.04 (0.40)</td>
<td>0.80 D</td>
<td>-0.04 (0.51)</td>
</tr>
<tr>
<td></td>
<td>Flat axis power (D):</td>
<td>0.36 (0.38)</td>
<td>2.76 D</td>
<td>-0.07 (1.89)</td>
<td>3.78 D</td>
<td>0.15 (1.93)</td>
</tr>
<tr>
<td></td>
<td>Astigmatism (D):</td>
<td>0.09 (0.62)</td>
<td>1.24 D</td>
<td>-0.13 (0.41)</td>
<td>0.82 D</td>
<td>-0.02 (0.56)</td>
</tr>
<tr>
<td></td>
<td>Flat axis power (D):</td>
<td>0.52 (1.65)</td>
<td>3.30 D</td>
<td>0.08 (1.56)</td>
<td>3.12 D</td>
<td>0.30 (2.03)</td>
</tr>
<tr>
<td></td>
<td>Astigmatism (D):</td>
<td>2° (2°)</td>
<td>4°</td>
<td>4° (9°)</td>
<td>18°</td>
<td>4° (5°)</td>
</tr>
<tr>
<td></td>
<td>Flat axis power (D):</td>
<td>5° (6°)</td>
<td>12°</td>
<td>4° (6°)</td>
<td>12°</td>
<td>9° (17°)</td>
</tr>
<tr>
<td></td>
<td>Flat axis power (D):</td>
<td>3° (4°)</td>
<td>8°</td>
<td>5° (10°)</td>
<td>20°</td>
<td>5° (6°)</td>
</tr>
<tr>
<td></td>
<td>Flat axis power (D):</td>
<td>12° (15°)</td>
<td>30°</td>
<td>10° (18°)</td>
<td>36°</td>
<td>13° (17°)</td>
</tr>
</tbody>
</table>

Coefficient of repeatability (COR) calculated as 2 × SD.

1.332 recommended by Zeiss, because in order to directly compare the two instruments we wanted to eliminate the bias of a non-uniform keratometric index.

All three captured images with the TMS-1 were processed. Absolute scale topographic maps were obtained for each eye, and the non-orthogonal simK readings (power and axis) were obtained.

STATISTICAL ANALYSIS

In order to assess observers variability the differences in measurements for each of the two instruments were calculated and plotted against the mean of the measurements, according to previously described methods. The deviation of each measurement from the mean of the three measurements of the same observer for each cornea was calculated for the intraobserver variation. The deviation scores were plotted against the mean of the three measurements for each subject. For the interobserver variation, the difference from the mean of the two observers was plotted against the mean of both observers’ measurements (six measurements). Statistical analysis for determining reproducibility also included calculation of the coefficient of repeatability (COR), which is defined as twice the standard deviation of the differences.

Results

NORMAL CORNEAS

Repeatability of keratometer on normal corneas (Table 1)

The coefficients of repeatability (COR) for all variables measured were comparable for examiners 1 and 2. The mean variance (deviation score or mean difference) for steep and flat axes power, and astigmatism magnitude was zero or very close to zero for both examiners, although one of them (observer 1) was much more experienced than the other in the use of this particular keratometer model. The obtained intraobserver COR for steep axis power measurement (0.22 D and 0.18 D for observers 1 and 2 respectively), flat axis power (0.18 D and 0.16 D), astigmatism (0.26 D and 0.20 D), steep axis angle (6° for both observers), and flat axis location (6° and 8°), are all considered to be at an acceptable level for clinical use.

The interobserver COR is slightly higher than the ones obtained with intraobserver measurements, but still within acceptable clinical levels (Table 1). The interobserver astigmatism COR was 0.28 D with only 6/66 measurements (6%) showing a deviation score of more than 0.25 D from the mean value, and no measurements (0/66) with a deviation of more than 0.50 D.
Repeatability of TMS-1 on normal corneas (Table 1)
Significant differences were found between the intraobserver COR for the two observers. Observer 1 (experienced with the use of TMS-1), has demonstrated consistently significantly lower coefficient values for all variables except steep axis location, compared with observer 2 who was a novice user of the instrument. Although for observer 1, COR reach clinically acceptable levels (0.30 D for steep axis power, 0.44 D for flat axis power, 0.40 D for astigmatism magnitude, 26° for steep axis location, 13° for flat axis location) these values are higher than those with the same observer using the keratometer. For the novice (observer 2) the obtained COR was above what is considered to be clinically acceptable in measuring steep axis power (1.08 D), flat axis power (1.96 D), astigmatism magnitude (1.62 D), steep axis (22°), or flat axis location (30°). These results highlight the superior repeatability performance of the keratometer compared with videokeratography, as well as the greater implication of the user’s experience on the reproducibility of TMS-1 compared with the keratometer. For the astigmatism measurement, observer 1 had 4/33 measurements (12%) with a deviation score > 0.25 D, but in only one of these measurements the deviation was in the range 0.50 D–1.00 D (Fig 1A). For observer 2, there were 5/33 measurements (15%) with a deviation score > 0.25 D, but for three of these five measurements the deviation was more than 1.00 D (Fig 1B). It is also obvious from the plot diagram that the high deviation scores were observed for corneas with higher astigmatism. The higher variability in observer 2’s readings, affects the interobserver variability as well. All COR variables are outside the clinically acceptable limits. This is particularly so for the measurement of flat axis power (1.82 D), and axes location (40° for the steep axis, 42° for the flat axis).

POST-PKP CORNEAS
Repeatability of keratometer on post-PKP corneas (Table 2)
Both observers (1 and 3), equally experienced with the use of keratometer, have shown significantly wide COR for all variables apart from axes location (Table 2). For observer 1, mean of differences, SD of differences, and COR for all measured variables are significantly higher than those observed for normal corneas (Table 2). This indicates that the repeatability of the keratometer decreases when studying highly astigmatic corneas. For observer 3, COR and SD are even greater than for observer 1, although observer 3 showed a better performance in identifying the axis location. Figures 2A and B show the plot diagrams for the repeatability in astigmatism measurements for examiners 1 and 3 respectively. Results on interobserver repeatability of the keratometer on post-PKP corneas are also shown in Table 2; the interobserver astigmatism measurement repeatability is represented graphically in Figure 2C.

Repeatability of TMS-1 on post-PKP corneas (Table 2)
Both observers 1 and 3 (both very experienced with the TMS-1) demonstrated poor repeatability in the highly astigmatic group. The poor performance is highlighted by the high values of mean variance, SD, and COR for both examiners in the measurement of steep axis power, flat axis power, astigmatism magnitude, flat axis angle (Table 2). Compared with the keratometer, the TMS-1 demonstrated a much higher range of variability in measuring post-PKP corneas. The repeatability of the TMS-1 on post-PKP corneas is worse than on normal corneas, as shown by the measurements of the same observer (observer 1) on normal corneas (Table 1) and post-PKP corneas (Table 2). Intra- and interobserver variability of TMS-1 on measuring astigmatism on post-PKP corneas is shown in Figures 3A–C.
Clinical evaluation of keratometry and computerised video-keratography

641

Discussion

The intra- and interobserver repeatability results of the present study on normal corneas have shown that the keratometer reproducibility is excellent for all measured variables, at very acceptable clinical levels and independent of the observer’s level of experience with the instrument. Compared with the TMS-1, the 10 SL/O keratometer has shown a superior repeatability in measuring normal corneas. On the other hand, TMS-1 repeatability was found in our study to be observer related. The novice (observer 2) demonstrated much wider variability on his measurements compared with the experienced examiner. No such difference between the two observers was noticed for the keratometer. In contrast, observers of similar experience produced similar results with the TMS-1 on post-PKP corneas. Intraobserver variation with the TMS-1 was also found to be astigmatism related, and increases with increasing astigmatism. In a previous study, a significant difference in the repeatability of the CAVK measurements between the two investigators was found and it was concluded that the examiner’s technique plays a role in the measurements. While some investigators have found no statistical differences between observers others, testing the interobserver reliability of three corneal topographers, suggested that for the TMS, operator reliability improves with training. The present study also revealed that on measuring highly astigmatic post-PKP corneas, the keratometer becomes less repeatable than on normal corneas. However, compared with the TMS-1, the keratometer achieved a much superior reproducibility. The repeatability of TMS on post-PKP corneas was poor (interobserver COR were 1.88 D for steep meridian power, 3.86 D for flat meridian power, 4.06 D for astigmatism, 34° for astigmatic axis location). The different performance between keratometry and TMS-1 on highly astigmatic postoperative corneas indicates that the two instruments should not be used interchangeably in these cases. No previous studies have vigorously assessed intra- and interobserver variability of video keratography and keratometry on highly astigmatic human corneas. Previous studies showing a decreased accuracy of CAVK on test surfaces steeper or flatter than the normal corneas or on radially aspheric test surfaces support the findings of the present study.

Accurate focusing is very important in obtaining reliable results with CAVK. The TMS-1 employs a patented crossed laser apex focusing system. Two low power HeNe laser sources are used to produce spots which must be criss crossed precisely 160 µm within the stroma to achieve proper focusing. However, the machine itself does not provide any record or verification of proper focusing. The alignment is also joystick operated with a provided record in terms of left/right and up/down alignment. For the TMS-1, a 500 µm defocusing, has been calculated to produce an error up to 2.5 D for spheres of 60 D. Hannush et al believe that focusing techniques played a significant role in the accuracy and precision differences among the instruments reported in their studies. In practice, it is true that because of the stromal scattering, blurring of the laser spots that the TMS-1 uses for focusing occurs, and this may complicate the accurate focusing. This phenomenon was observed throughout this study, but for the experienced observer it did not seem to represent a serious problem. However, as the focusing system of TMS-1 is relatively subjective, dependent on the operator’s judgment, it may well have contributed to the worse results seen with the novice observer. Mandell also found a considerable interobserver variation in the effort expended to achieve a perfect focus with the TMS-1. According to the same author, the EyeSys system has a less critical focusing problem as a result of the greater distance between instrument and target. The TMS-1 projects 25 or 32 rings of cool green light onto the cornea from a
cylindrical Placido through a light cone at a distance of 32 mm from the 12th ring, and this shorter working distance may be prone to larger measurement errors if the instrument is defocused. However, McCarey et al. using the EyeSys system, found that variations in focusing beyond 0.25 mm from the centre result in unreliable data, and results vary considerably on defocusing beyond 1 mm. Furthermore, similar differences in results between keratometry and a device using autofocus feedback mechanism have been obtained. 16 On the other hand, for the keratometer, incorrect focusing of the eyepieces of the instrument produce inaccuracies. This can produce an error in the radius reading as much as 0.4 mm. 8 Additional inaccuracies of a mechanical nature (backlash) can be caused by the possible amount of freedom of movement of the control knob before it operates the movable mechanism of the instrument. 16 Although “one position” keratometers in particular, tend to be quite incorrect for one of the two meridians because only one meridian can be in focus in toroidal corneas, the 10 SL/O Zeiss keratometer requires separate focusing for steep and flat meridians and therefore does not suffer from this disadvantage. The placement of the single cross centrally within the double cross as the instrument requires, is much less skill and experience related than the focusing of the TMS-1. Despite the fundamental limitations of the keratoscopic design, as errors can be introduced owing to poor focus of different rings 17 if these are not on the same plane, TMS-1 manufacturers claim that the patented target eliminates this problem, by using a Placido cone rather than flat target. Tsilimbaris et al., 9 using a system with a different keratoscope (EyeSys), suggested the observed differences in astigmatism were the result of the problem in focusing of both principal meridians at the same time with the EyeSys in highly astigmatic corneas. Legeais et al. 5 also suggested that the accuracy of the TMS could be improved by increasing the depth of field, and to achieve that a stronger illumination source with a higher sensitivity charged coupled device camera could be used.

Correct alignment is also very important with videokeratoscopy. In order for the CAVK to fulfil many of the assumptions made by the reconstructive algorithms, the cornea must be correctly positioned. 10 Small errors in alignment can result in an irregular or asymmetric topographic reconstruction. In one commercially available instrument (EyeSys) poor fixation has been shown to produce a pattern of pseudokeratoconus; the increase in relative steepness was statistically significant at 5 degrees of deviation. 20 Therefore, errors in focusing and misalignment may have contributed to the novice observer variation compared with the experienced observer, as investigators not very familiar with the joystick alignment of the TMS-1 may be prone to slight decentrations and defocusing.

In summary, the findings of the present study indicate that although keratometer variability for normal corneas is excellent and not affected by instrument type, the TMS-1 videokeratoscope repeatability is observer related and can be improved with better training. Intraobserver variation with the TMS-1 is also astigmatism related, and increases with increasing astigmatism. On the basis of the above findings it is finally concluded that the two instruments demonstrate clinically significant differences both on normal and astigmatic corneas and therefore should not be used interchangeably, especially on highly astigmatic corneas. For the TMS-1, users with the same level of experience with the instrument should be employed in clinical or experimental studies.

This study was supported in part by a grant from the Hellenic Scholarships Foundation (IKY).

References