Repeatability and reproducibility of manifest refraction

Suphi Taneri, MD, FEBOS-CR, Samuel Arba-Mosquera, PhD, Anika Rost, MSc, Saskia Kießler, Dipl-Ing, H. Burkhard Dick, MD, FEBOS-CR

Purpose: To evaluate the intraexaminer repeatability and the interobserver reproducibility of manifest refraction.

Setting: Tertiary referral center.

Design: Retrospective study.

Methods: Patients attending at least 2 preoperative refractions before undergoing subsequent refractive surgery were included. All manifest refractions were performed by 1 or 2 different optometrists using an automated phoropter according to a standard protocol. The first manifest refraction was performed after obtaining automated refraction and measuring the spectacles of the patient. The second refraction was typically refined from the first also considering wavefront refraction and tomography/topography.

Results: The latest 2 manifest refractions of 1000 eyes obtained at 2 separate visits showed a mean pairwise absolute difference of 0.16 ± 0.19 diopter (D) (range 0 to 1.38 D) in spherical equivalent (SE). This SD was better than 0.25 D (the minimum measurement increment of refraction itself). The 95% limit of agreement (LoA) was within 0.50 D for sphere, cylinder, and SE. The SD of the astigmatism axis was approximately 10 degrees, and the 95% LoA was within 22 degrees (the difference in axis decreasing significantly with the measured cylinder magnitude). The SD for corrected distance visual acuity (CDVA) was half a Snellen line and the 95% LoA was within 1.5 lines (with increasing deviation with worse vision). There were no clinically meaningful differences in reproducibility (2 optometrists) compared with repeatability (same optometrist) in sphere, axis, and CDVA.

Conclusions: Reproducibility was 0.16 D irrespective whether refractions were performed by 1 or 2 different optometrists. Obtaining multiple refractions preoperatively might increase the predictability of surgery and decrease the enhancement rate.

Laser refractive surgery has grown in complexity through the years. Initial treatments were merely based on manifest refraction. By contrast, current state-of-the-art treatment planning includes consideration of topographic, tomographic, and pachymetric findings, axial length, epithelial mapping, wavefront measurements, and biomechanical aspects. This information is used both in patient selection and for refined treatment planning. However, the most influential examination, measuring manifest refraction, is essentially still performed in much the same manner as 30 years ago. All other above-mentioned measurements eventually aid in advancing the treatment plan by more accurately describing the refractive defect.

The most accepted review determined the minimum uncertainties (because other factors such as unwanted accommodation might also be involved) and showed that the standard uncertainty (±1 SD) is about 0.3 diopter (D) in refractive error measurement and about 0.04 in logarithm of the minimum angle of resolution (logMAR) visual acuity.1 The normally quoted expanded uncertainty, which provides a 95% confidence level, would then be 0.6 D in refractive measurement and 0.08 in logMAR acuity. The ISO issued code 13666 for ophthalmic optics and spectacle lenses and provided similar uncertainties.2 Aberrometric refraction has been proposed as an alternative to overcome some of the problems associated with manifest refraction, but the repeatability remains at the 0.25 D level.3 Other refraction methods have also been proposed recently to avoid some of the issues.4 We wanted to study the repeatability and reproducibility of manifest refraction and corrected distance visual acuity (CDVA) in normal preoperative patients following a systematic protocol in standardized conditions in our refractive surgery practice.

Submitted: May 7, 2020 | Final revision submitted: June 22, 2020 | Accepted: June 25, 2020

From the Center for Refractive Surgery, Eye Department at St. Francis Hospital (Taneri, Rost, Kießler), Muenster, Ruhr-University, University Eye Hospital (Taneri, Dick), Bochum, and SCHWIND eye-tech-solutions (Arba-Mosquera), Kleinostheim, Germany.

Presented in part at the German Society of Cataract and Refractive Surgery Annual Meeting, Mainz, Germany, February 2020.

Caroline Hansson, MSc, and Katja Rickels, BSc, assisted in data acquisition.

Corresponding author: Suphi Taneri, MD, FEBOS-CR, Center for Refractive Surgery, Hohenzollernring 70, 48145 Muenster, Germany.

Email: taneri@refraktives-zentrum.de.

Copyright © 2020 Published by Wolters Kluwer on behalf of ASCRS and ESCR.

https://doi.org/10.1097/j.jcrs.0000000000000343

0886-3350/$ - see frontmatter
METHODS

Patient Population and Examinations
The study adhered to the tenets to the Declaration of Helsinki, and informed consent was obtained from all participants for the anonymized evaluation of their data. Because of the retrospective nature of this study ethics, committee approval was not needed.

Electronic database was searched for the last 1000 consecutive myopic or myopic astigmatic eyes in which at least 2 manifest refractions were performed during the preoperative period. The time interval between the 2 visits was at least 48 hours and not more than 3 months. Refractions were performed by 1 of 4 experienced optometrists.

Exclusion criteria were medical conditions or medication potentially affecting refractive power or visual acuity, including less than perfectly controlled diabetes mellitus, a history or slitlamp signs of ocular trauma or ocular surgery, and systemic use of corticosteroid, antimitabolite, or immunosuppressant agents. Patients with a history, signs, or symptoms of recurrent basement membrane dystrophy or recurrent erosion syndrome were also excluded.

Standard Workflow
The authors’ standard workflow of adult patients seeking vision correction comprises at least 2 evaluations on 2 different days. Soft contact lens wear must be discontinued for at least 2 weeks and rigid contact lens wear for at least 4 weeks before the initial examination. Special emphasis is paid to optimize the ocular surface. The manifest refraction is performed at the initial visit and repeated at least 1 separate visit preoperatively to determine the final refraction to plan the treatment.

Manifest Refraction Protocol Standardized Conditions. The authors’ process of measuring refractive error and visual acuity is based on a standardized protocol devised by 1 of the authors specifying the ISO recommendations for their setting. All 4 optometrists were trained in applying this protocol. In their center, the same room (lane) was routinely used for performing preoperative manifest refraction, with a constant distance of 6 m from phoroptor (Visutron, Möller-Wedel Optical Gmbh) to the optotype screen and constant lighting conditions (no natural daylight). Randomly generated normed optotypes (Landolt rings, no numbers) with a constant contrast are presented on a computer screen (Multivius optotype, bon Optic). Cross-cylinder method was used. The process was concluded with fine-tuning the cylinder magnitude and axis displaying a “sun dial” on the screen to check for homogeneous vision (Figure 1).

First Evaluation. During the first evaluation, refraction history and spectacle prescription were recorded using an automated lensmeter (LM-990A; NIDEK Co., Ltd.). Automated refraction without pharmacologic pupil dilation was obtained averaging 5 single measurements (Canon R-F10, Canon, Inc.). Topography and tomography (Pentacam AXI, OCULUS Optikgeräte GmbH, and Orbscan IiX, Bausch & Lomb, Inc., respectively) were performed twice to better determine corneal astigmatism magnitude and axis. Manifest refraction along with CDVA was performed by protocol, starting with spectacles.

Second Evaluation. At a different day, taking into account all details of the first evaluation and the wavefront refraction (Zywave II, Bausch & Lomb, Inc.), a second manifest refraction (starting by first refraction) along with CDVA, cycloplegic autorefraction (under the use of tropicamide), and cycloplegic refraction along with cycloplegic CDVA were performed.

Third Evaluation (if discrepancy between first and second evaluations). In cases where first and second manifest refractions or manifest refraction and any other examinations (tomography, topography, and cycloplegic autorefraction) reveal significant discrepancies (as evaluated by S.T.), then, a third evaluation was performed. This evaluation included at least manifest refraction along with CDVA; other examinations were repeated as deemed necessary and/or surgeon’s (confirmatory) manifest refraction was performed. If the discrepancy was not resolved, further evaluations were scheduled to allow more time for optimizing the ocular surface, which might still be affected by dryness or contact lens warpage. However, only the latest 2 refractions (after stabilization) were evaluated in this study.

Data Analysis
All data were analyzed using Excel (Microsoft Corp.). The 2 latest manifest refractions and associated CDVA of 1000 consecutive eyes were compared. Stratification of eyes was performed with both refractions performed by either the same optometrist (group 1) or by 2 different optometrists (group 2). Bland-Altman and Box-and-Whisker plots were used to present the findings (both arithmetically and in absolute value). The standard graphs for reporting outcomes in refractive surgery have been adapted to display the differences.

The significance of the differences was evaluated considering a metric distributed approximately as t with N—degrees of freedom, where N is the size of the sample considered as number of patients (and not number of eyes). Paired t tests or analysis of variance tests were used to determine statistically significant changes. A P value less than 0.05 was considered statistically significant. Only preoperative data are reported in this study.

RESULTS

The most recent 1000 consecutive eyes of 510 patients fulfilling the inclusion criteria were included in the analyses: 497 eyes were refracted by the same optometrist (group 1) and 503 eyes were refracted by 2 different optometrists (group 2). Demographics are presented in Table 1.

Results of the Whole Cohort

Box-and-Whisker Plots The differences for the whole cohort between 2 refractions in sphere, cylinder, spherical equivalent (SE), and CDVA are displayed in Figure 2 and in astigmatism axis in Figure 3. Figure 2, a shows the arithmetic values, while Figure 2, b shows the absolute values.
The mean difference in sphere was 0.03 ± 0.24 D and 0.15 ± 0.19 D in absolute value. The mean difference in cylinder was 0.02 ± 0.21 D and 0.12 ± 0.18 D in absolute value. The mean difference in SE was 0.05 ± 0.25 D and 0.16 ± 0.19 D in absolute value. The mean difference in CDVA was 0.02 ± 0.07 logMAR and 0.05 ± 0.06 logMAR in absolute value. The mean difference in astigmatic axis was 0.1 ± 11.1 degrees and 5.1 ± 9.8 degrees in absolute value.

**Bland-Altman Plots**
The differences between 2 refractions for the whole cohort are displayed as Bland-Altman plots in Figure 4 for sphere and SE, in Figure 5 for astigmatism magnitude and axis, and in Figure 6 for CDVA. The difference in sphere showed no correlation with the measured spherical magnitude, but the difference in absolute value increased significantly (P < .05) with the measured spherical magnitude, and the 95% limit of agreement (LoA) ranged from −0.43 D to +0.50 D difference (within 0.50 D in absolute value) (Figure 4, a and b). The difference in SE increased significantly (P < .05) with the measured SE magnitude, and the 95% LOA ranged from −0.44 D to +0.53 D difference (within 0.50 D in absolute value) (Figure 4, c and d). The difference in cylinder showed no correlation with the measured cylinder magnitude, but the difference in absolute value increased significantly (P < .05) with the measured cylinder magnitude, and the 95% LOA ranged from −0.39 D to +0.44 D difference (within 0.50 D in absolute value) (Figure 5, a and b). The difference in axis showed no correlation with the measured axis, and the 95% LOA ranged from −22 degrees to +22 degrees difference (within 20 degrees in absolute value) (Figure 5, c and d). However, the difference in axis decreased significantly (P < .05)

<table>
<thead>
<tr>
<th>Table 1. Demographics.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parameter</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Eyes/patients (n)</td>
</tr>
<tr>
<td>Age (y)</td>
</tr>
<tr>
<td>Sex (F/M)</td>
</tr>
<tr>
<td>Spherical equivalent (D)</td>
</tr>
<tr>
<td>Mean ± SD</td>
</tr>
<tr>
<td>Range</td>
</tr>
<tr>
<td>Cylinder (D)</td>
</tr>
</tbody>
</table>
with the measured cylinder magnitude (Figure 5, e and f). The difference in CDVA increased significantly ($P < .05$) for worse CDVA values, and the 95% LOA ranged from $-0.12$ logMAR to $+0.17$ logMAR difference (within $0.12$ logMAR in absolute value) (Figure 6).

**Same vs Different Optometrist Comparison**

The comparative analyses of measurements by the same vs 2 different optometrists are displayed in Figure 7 for CDVA, SE, cylinder, and astigmatism axis. There was a statistically significant but not clinically meaningful difference.
In brief, in our study, the SD of refraction was better than 0.25 D (the minimum measurement increment of refraction itself), and the 95% LoA was within 0.50 D for sphere, cylinder, and SE. The SD of the astigmatism axis was approximately 10 degrees, and the 95% LoA was within 22 degrees (the difference in axis decreasing significantly with the measured cylinder magnitude). The SD for CDVA was half a Snellen line, and the 95% LoA was within 1.5 lines (with increasing deviation with worse vision). There were minor differences in reproducibility (2 optometrists) compared with repeatability (same optometrist) in sphere, axis, and CDVA. However, these differences were not clinically meaningful despite reaching statistical significance. 

Derived from our results and assuming a 0.25 D threshold, the following may be concluded: For sphere, it seems that the absolute difference between 2 measurements is approximately 2.6% of the actual value (beyond −13 D, it transcends 0.25 D). For cylinder, it seems that the absolute difference between 2 measurements is approximately 7.9% of the actual value (beyond 7 D, it transcends 0.25 D). For SE, it seems that the absolute difference between 2 measurements is approximately 2.7% of the actual value (beyond −12 D, it transcends 0.25 D). For CDVA, it seems that the absolute difference between 2 measurements reaches 1 Snellen line at 0.4 logMAR (20/50) CDVA.

Determination of the manifest refraction as accurately as possible is of utmost importance for several reasons. Obviously, despite substantial advancements in objective evaluation of the individual eye and its refractive error, subjective manifest refraction is still the most influential parameter in every refractive treatment performed today, except for refractive lens exchange. In addition, postoperatively, manifest refraction serves as feedback signal to monitor and refine the surgical nomogram.

In this study, we tried to examine repeatability and reproducibility of manifest refraction in a specialized refractive clinic optimized for exactly that task, measuring refraction. The conditions described in this study might be as good as possible in a real-world setting because only cooperative and highly motivated adult patients seeking refractive surgery, and only healthy eyes without previous surgery were examined. Moreover, examiners had access to all state-of-the-art objective examinations including automated refraction (both with and without cycloplegia), topography/tomography, and wavefront aberration refraction; in addition, environmental parameters were kept as constant as possible. Besides, in our protocol, special attention is paid to optimizing the ocular surface before obtaining manifest refraction that might be used for planning a refractive treatment as a poor ocular surface would compromise the complete evaluation including refraction, visual acuity, keratometer readings, topography/tomography, and pachymetry.

Therefore, only the latest 2 refraction measurements with an acceptable ocular surface were evaluated in this study (if more than the latest 2 refractions had been evaluated, the differences had been greater—data not shown). However, measuring refraction remains challenging even under these “ideal” circumstances because the refractive error is fluctuating. Factors affecting the refractive status include accommodation, stress level, fatigue, and ambient illumination conditions, apart from long-term drifts.

The simultaneous change of various parameters makes it very difficult to determine the isolated effects of the
different individual parameters, because the signal-to-noise ratio is reduced. In principle, there are 2 approaches to account for that. One approach is to largely increase the sample size, and the other one is to perform controlled studies, aiming to isolate the effect of individual parameters by having a cohort in which all parameters are identical, except for the one under test.

We followed a combined approach in this study. We performed a careful study for which as many parameters as possible were kept constant including same visual lane, fixed distance of 6 m from subject to the optotype screen, constant lighting conditions (no natural daylight), randomly generated normed optotypes (no numbers) with a constant optotype contrast, a standardized refraction protocol, and only healthy myopic or myopic astigmatic eyes. We also included a large sample size (1000 consecutive eyes). The 2 different refractions considered were taken on different days (minimum 48 hours, maximum 3 months apart) by either the same or 2 different examiners (out of a team of 4 experienced optometrists).

Comparing both refractions (regardless of these being measured by the same or 2 different optometrists), overall mean differences were close to zero, and the SD of the differences were below 0.25 D for refraction (sphere, cylinder, and SE both arithmetically and in absolute value), below 1 line for CDVA, and below 15 degrees for astigmatism axis. Of interest, the difference in axis decreased significantly with the measured cylinder magnitude. We assume that the greater influence on visual quality of an axis change in higher astigmatism values might be the main reason for a more reliable measurement. The difference in CDVA significantly increased with worse measured CDVA values, and the 95% LoA remained within 1.5 Snellen lines.

The comparative analyses for same vs different optometrists showed that different optometrists obtain a statistically but not clinically significantly higher variability than a single optometrist for sphere, astigmatism axis, and CDVA. Of note, 1 amblyopic eye in our cohort showed deviations that were clearly out of the normal distribution (a difference of 1.75 D in sphere and cylinder magnitude). This might be explained by more intense pushing of the cylinder at the second refraction after the wavefront refraction had revealed a high cylinder.

We have methodically presented the findings arithmetically and in absolute values. The arithmetical comparison reveals systematic differences between both refractions and/or different optometrists, which were not observed. By contrast, the analysis of the differences in absolute values corresponds better to a measurement of the variability because it is irrelevant which refraction was taken first. Overall, the findings are comparable or slightly better than previously reported values, which might partly be due to our patient selection and the additional information from objective examinations in our setting.

Regarding the study design, 2 aspects merit clarification. First, we believe that the retrospective nature of this chart review better represents the real values in our setting (a refractive surgery clinic) than a prospective study could because it avoids a significant potential bias, namely that the examiners would be aware of participating in a study. Second, it might be interpreted as a major limitation that...
the 2 separate refractions were not truly independent, and further to that, in case of large disparities, a third refraction measurement was appointed (but only the latest 2 were evaluated here). However, the aim of this study was not to compare the agreement of 2 refractions that were taken without any further information in all comers. Rather, our aim was to evaluate refraction in refractive patients under real-world conditions, which include a standardized approach under constant conditions, which naturally reduce the variability of the measurements compared with truly unbiased independent refractions. So, care must be taken not to generalize our results to less-controlled settings such as clinics with other subspecialties.

Limitations of our study include that the time interval between the 2 manifest refractions was not identical for all patients and that the examinations were not performed at the same time of day. Because the actual manifest refraction fluctuates within a small range diurnally and might also drift over time, some pseudo errors might have been introduced. Furthermore, this is a single-center study, and a multicenter trial is desirable. However, the setting, the process of measurement, and the cohort compositions would most likely be different making comparisons difficult. Finally, all 4 examiners in our study had several years of experience. So, we might not generalize our results to optometrists in training. Similarly, because we analyzed only adult patients, our results might not be applicable in minors. Because we have not correlated the 2 refraction measurements with our treatment results, it is beyond the scope of this article to decide whether the first or the second measurement is in better agreement with the surgical outcome. However, when actually planning a treatment in our clinical routine, we look at all the individual examinations (including corneal astigmatism in topography/tomography, wavefront refractions, etc.) and come up with a treatment refraction that is a weighted average influenced by the surgeon’s discretion. The resulting treatment refraction might, therefore, be different than both refraction measurements. In addition, our guided standardized protocol that requires at least 2 measurements at 2 different days is based on theoretical considerations and clinical experience (only). It has evolved over time and has been constant for the past 5 years. However, we have not compared this approach to alternatives such as routinely performing 3 refractions or to performing multiple refractions on a single day.

Our results suggest that, using a guided standardized protocol in a controlled environment, the repeatability and reproducibility of manifest refraction in a select group of eyes (of refractive patients) is better than the typical step used for manifest refraction (0.25 D). There were no clinically relevant differences in manifest refraction when comparing the results obtained by the same vs 2 different optometrists, which might be interpreted as a benefit of standardizing the process of refraction measurement. Obtaining multiple refractions preoperatively might increase the predictability of refractive surgery and, thus, decrease the enhancement rate.

WHAT WAS KNOWN

- Refraction is the most important parameter in laser refractive surgery.
- Refraction is influenced by pupil size, accommodation, and hormonal level among other factors.

WHAT THIS PAPER ADDS

- Using a standard protocol for refraction measurement in a very controlled setting, there was no clinically meaningful difference whether repeat refraction was performed by 1 or 2 different examiners.
- Mean difference SD was below 0.25 diopter.

REFERENCES

19. Raasch TW, Schechtman KB, Davis LJ, Zadinik K; CLEK Study Group. Collaborative Longitudinal Evaluation of Keratoconus Study. Repeatability of


**Disclosures:** Dr. Taneri is consultant to Bausch & Lomb, Inc. and Carl Zeiss Meditech AG. None of the other authors have a financial or proprietary interest in any material or method mentioned.