Reproducibility of manifest refraction between surgeons and optometrists in a clinical refractive surgery practice

Dan Z. Reinstein, MD, MA(Cantab), FRCSC, DABO, FRCOphth, FEBO,
Timothy E. Yap, BMedSci (Hons), MA, Glenn I. Carp, MB BCh, FC Ophth (SA),
Timothy J. Archer, MA(Oxon), DipCompSci(Cantab), Marine Gobbe, PhD, MSTOptom,
and the London Vision Clinic optometric group

PURPOSE: To measure and compare the interobserver reproducibility of manifest refraction according to a standardized protocol for normal preoperative patients in a refractive surgery practice.

SETTING: Private clinic, London, United Kingdom.

DESIGN: Retrospective case series.

METHODS: This retrospective study comprised patients attending 2 preoperative refractions before laser vision correction. The first manifest refraction was performed by 1 of 7 optometrists and the second manifest refraction by 1 of 2 surgeons, all trained using a standard manifest refraction protocol. Spherocylindrical data were converted into power vectors for analysis. The dioptric power differences between observers were calculated and analyzed.

RESULTS: One thousand nine hundred twenty-two consecutive eyes were stratified into a myopia group and a hyperopia group and then further stratified by each surgeon–optometrist combination. The mean surgeon–optometrist dioptric power difference was 0.21 diopter (D) (range 0.15 to 0.32 D). The mean difference in spherical equivalent refraction was 0.03 D, with 95% of all refractions within $\pm$0.44 D for all optometrist–surgeon combinations. The severity of myopic or hyperopic ametropia did not affect the interobserver reproducibility of the manifest refraction.

CONCLUSIONS: There was close agreement in refraction between surgeons and optometrists using a standard manifest refraction protocol of less than 0.25 D. This degree of interobserver repeatability is similar to that in intraobserver repeatability studies published to date and may represent the value of training and the use of a standard manifest refraction protocol between refraction observers in a refractive surgery practice involving co-management between surgeons and optometrists.

Financial Disclosure: Dr. Reinstein is a consultant to Carl Zeiss Meditec AG and has a proprietary interest in the Artemis technology, Arcscan, Inc., through patents administered by the Cornell Center for Technology Enterprise and Commercialization. No other author has a financial or proprietary interest in any material or method mentioned.

J Cataract Refract Surg 2014; 40:450–459 © 2014 ASCRS and ESCRS

Supplemental material available at www.jcrsjournal.org.

Measuring refractive error is a process that involves a subjective component that is open to interobserver and intraobserver variability. It is important to know the size of this variability to qualify comparisons between measurements and detect significant changes. The accuracy of refractive surgery depends on nomogram optimization, which in turn depends on the accuracy of refraction (the agreement between the measured actual refraction) and the reproducibility of refraction (the concordance between observers) before surgery for surgical planning as well as after surgery for outcomes feedback.[1–6] The co-management of refractive surgery care between optometrists and surgeons is a well-established and common model in practice, with optometrists working in a surgeon's practice or working outside the practice. Therefore, in any refractive surgery clinical setting, it is important that the accuracy and reproducibility of manifest refraction

Published by Elsevier Inc.
0886-3350/$ - see front matter
http://dx.doi.org/10.1016/j.jcrs.2013.08.053

within and between clinicians (surgeons and optometrists) is optimized to optimize clinical outcomes.

Previous studies measured the repeatability of subjective manifest refraction with the same examiner (intraobserver)[7,8] or the reproducibility between different examiners trained to give standardized responses who perform refraction in patients (interobserver).[9,10] Repeatability studies in controlled study settings report a median difference of 0.20 diopter (D) with a difference of 0.62 D or less in 95% of cases,[7] while reproducibility of the spherical equivalent (SE) refraction has been reported to be within ±0.25 D in 81% of eyes.[9] To our knowledge, a retrospectively analyzed reproducibility study between multiple observers in a large patient population in a routine clinical refractive surgical practice has not been published.

The aim of the present study was to evaluate the reproducibility of manifest refraction in a large patient population between each of 2 surgeons and 7 optometrists, all trained and working under a standardized manifest refraction protocol in a single refractive surgery practice. The study also sought to determine whether the magnitude of myopia or hyperopia influences the reproducibility of manifest refraction and whether refractive surgery outcomes are influenced by systematic differences that may have been discovered between observers.

PATIENTS AND METHODS

This was a retrospective study of consecutive patients who had laser in situ keratomileusis (LASIK) at the London Vision Clinic, London, United Kingdom, between February 2009 and January 2012. The preoperative assessment included an initial visit performed by 1 of 7 in-house optometrists and a second visit performed by 1 of 2 surgeons (D.Z.R [Surgeon A] and G.I.C. [Surgeon B]).

Submitted: February 4, 2013.
Final revision submitted: June 28, 2013.
Accepted: August 15, 2013.

From the London Vision Clinic (Reinstein, Yap, Carp, Archer, Gobbe), London, United Kingdom; Department of Ophthalmology (Reinstein), Columbia University Medical Center, New York, New York, USA; Centre Hospitalier National d'Ophtalmologie des Quinze-Vingts (Reinstein), Paris, France.

The following optometrists from the London Vision optometric group participated in the study: Alexandra Lyons, BSc(Hons), MCOptom, Brendan Duane, BSc(Optom), Connie McEvoy, BSc(Hons), MCOptom, Desiree Todd, BSc(Hons), Emma Brandon, BSc(Hons), MCOptom, MSc, Sharon Ritchie, BSc(Hons), MCOptom, Vimal Piparia, BSc(Optom).

Corresponding author: Dan Z. Reinstein, MD, MA(Cantab), FRCSC, FRCOphth, London Vision Clinic, 138 Harley Street, London W1G 7LA, United Kingdom. E-mail: dzr@londonvisionclinic.com.

Inclusion criteria were that the patients be medically suitable for LASIK; have no previous ocular, eyelid, or orbital surgery; have a corrected distance visual acuity (CDVA) of 20/20 or better; were between 18 years and 60 years old; did not wear rigid contact lenses for at least 6 months or soft contact lenses for at least 1 month before the first manifest refraction; and did not wear contact lenses between the first refraction and second refraction. In addition, the time between the 2 manifest refractions could be no more than 60 days.

Refraction Procedure

The first appointment was conducted by 1 of 7 in-house optometrists. Before the optometrist's eye examination, if the patient was wearing spectacles, the refraction of these was measured with a lensometer (Xeed-1, Canon, Inc.). A refraction was also obtained from an undilated Hartman-Shack wavefront aberration–supported cornea ablation aberrometry scan (WASCA, Carl Zeiss Meditec AG). The optometrist then performed a manifest refraction using a phoropter using a cross-cylinder technique for cylinder refinement and the CSV-1000 Early Treatment of Diabetic Retinopathy Study chart (Vector Vision). Lighting conditions were controlled with ambient light set to mesopic levels (3 lux). The aberrometric refraction was taken as the starting point for the manifest refraction. The manifest refraction was performed based on a standardized protocol devised by 1 of the authors (D.Z.R.) and published previously,[11] and all optometrists and surgeons were formally trained in applying this protocol (see Appendix A, available at http://jcrsjournal.org). After the manifest refraction, 1 drop of tropicamide 1.0% was used to induce cycloplegia and the optometrist performed a cycloplegic refraction. The patient returned at least 1 day later for a consultation with the surgeon, who performed a second manifest refraction. Before starting the manifest refraction, the surgeon reviewed the 4 previous refractions (spectacles refraction, undilated aberrometric refraction, optometrist manifest refraction, and optometrist cycloplegic refraction). The surgeon then performed a manifest refraction using the same refraction protocol, starting from the optometrist's manifest refraction or the optometrist's cycloplegic sphere if this showed a latent component relative to the manifest refraction.

The clinic's routine protocol was to see patients 1 day, 1 month, 3 months, 6 months, and 12 months after surgery. A manifest refraction was performed at each visit by 1 of the 7 optometrists. The postoperative data were used to assess the clinical significance of differences found between surgeons.

Statistical Analysis

Spherocylindrical data from the optometrist and surgeon manifest refractions were transformed into power vectors as described by Thibos and Horner[12] The M, J0, and J45 components (SE, x-coordinate, and y-coordinate of the astigmatism vector, respectively) were calculated along with the resultant total dioptric power as follows:

$$\sqrt{(M^2 + J0^2 + J45^2)}$$

The difference in total dioptric power was calculated as follows:

$$\sqrt{\left[(\Delta M)^2 + (\Delta J0)^2 + (\Delta J45)^2\right]}$$

where Δ is the difference between the surgeon refraction and optometrist refraction.

Eyes were divided into 2 groups (myopia and hyperopia) according to refractive error, with myopia defined as sphere of 0.00 D or less according to the surgeon's refraction. These groups were further subdivided for each of the 14 possible optometrist–surgeon combinations.

Descriptive statistics (mean, standard deviation [SD], minimum, maximum) were calculated, the mean and SD were plotted, and 1-way analysis of variance (ANOVA) was performed for the difference in sphere, cylinder magnitude, and total dioptric power for each optometrist–surgeon combination in the myopia group and hyperopia group. The mean difference and 95% limits of agreement (LoA) (1.96 × SD) were calculated for the SE refraction (M), J0, J45, and cylinder vector difference for each optometrist–surgeon combination in the myopia group and hyperopia group. The intraclass correlation coefficient (ICC) was calculated for the total dioptric power difference between the surgeon and the optometrist for each surgeon in the myopia group and hyperopia group. Tukey honestly-significant-difference (HSD) post hoc tests were performed to identify the optometrist pairs for which there was a statistically significant difference. An F test was performed to test for a difference in SDs. The percentage of eyes in which the total dioptric power difference was 0.25 D or less or 0.50 D or less was calculated for each optometrist–surgeon combination in the myopia group and hyperopia group. The mean accuracy of the refractive outcome in the myopia group and hyperopia group was calculated for each surgeon based on the last follow-up data.

A Bland-Altman chart was plotted for SE refraction in all eyes to determine whether the difference in manifest refractions between the optometrist and the surgeon was affected by the magnitude of refractive error. Scatterplots were created to determine whether the difference in SE refraction between the optometrist and the surgeon was affected by the time between visits in the myopia group and hyperopia group.

Excel 2010 software (Microsoft Corp.) was used for data entry and statistical analysis except for the 1-way ANOVA and Tukey HSD post hoc tests, which were calculated using SPSS software (version 20, International Business Machines Corp.). A *P* value of less than 0.05 was defined as statistically significant.

RESULTS

After the inclusion criteria were applied to consecutively treated patients, a qualifying optometrist–surgeon pair of manifest refractions was available for 1922 eyes, including 954 myopic eyes and 968 hyperopic eyes. Table 1 shows the demographic data of the study population grouped by surgeon; the groups were well matched. Appendix B and Appendix C (available at http://jcrsjournal.org) show the demographic data in the myopia group and hyperopia group, respectively, for all surgeon–optometrist combinations.

Table 2 shows the ICC for the total dioptric power difference in the myopia group and hyperopia group between each surgeon and the optometrists. The ICCs were all very close to 0, which indicates that each eye could be considered independent and therefore both eyes of all patients could be included in the analysis.

Figure 1 shows the mean and SD for the difference in total dioptric power, sphere, and cylinder magnitude in the myopia group and hyperopia group for each optometrist–surgeon combination. Table 3 shows the percentage of eyes in which the total dioptric power difference in the myopia group and hyperopia group was 0.25 D or less or 0.50 D or less for each optometrist–surgeon combination. Overall, including all optometrist–surgeon combinations, the mean total dioptric power difference was 0.21 D (median

Table 1. Patient demographics.

	Myopia Group		Hyperopia Group	
Parameter	Surgeon A	Surgeon B	Surgeon A	Surgeon B
Age (y)				
Mean ± SD	42 ± 10	41 ± 11	53 ± 7	52 ± 8
Range	19, 61	19, 60	18, 61	18, 65
Male sex (%)	56	62	36	46
Preop sphere (D)				
Mean ± SD	−2.98 ± 2.17	−2.86 ± 2.33	+1.74 ± 1.21	+2.05 ± 1.57*
Range	0.00, −13.00	0.00, −12.25	+0.25, +8.75	+0.25, +8.50
Preop cylinder (D)				
Mean ± SD	−0.96 ± 0.78	−0.92 ± 0.74	−0.73 ± 0.79	−0.76 ± 0.84
Range	0.00, −4.00	0.00, −5.25	0.00, −5.25	0.00, −6.00
Time between refractions (days)				
Mean ± SD	15 ± 16	13 ± 14*	16 ± 16	14 ± 16*
Range	1, 61	1, 60	1, 62	1, 62
Eyes/patients (n)	487/260	467/244	492/271	476/257

*$P < .05$ between surgeons

Table 2. Intraclass correlation coefficient for total dioptric power difference between surgeon and optometrist.

Group/Surgeon	ICC
Myopia	
Surgeon A	0.0077
Surgeon B	0.0160
Hyperopia	
Surgeon A	−0.0045
Surgeon B	0.0292

ICC = intraclass correlation coefficient

0.19 D); it was less than 0.25 D in 1219 eyes (63%) and less than 0.50 D in 1798 eyes (94%). The mean difference in SE refraction was 0.03 D (95% LoA, ±0.44 D). Table 4 shows the mean difference and 95% LoA for M, J0, J45, and the vector difference of cylinder in the myopia group and hyperopia group for each optometrist–surgeon combination.

Table 5 shows the optometrist–surgeon combinations for which a statistically significant difference was identified in the 1-way ANOVA for the total dioptric power, sphere, and cylinder magnitude in the myopia group and hyperopia group. There was no statistically significant difference in mean values between optometrists in total dioptric power compared with either surgeon in the myopia group and the hyperopia group. There was also no statistically significant difference in mean cylinder magnitude values in the hyperopia group between optometrists compared with Surgeon B. There was a small but statistically significant difference between 1 and 7 optometrist pairs in the other subgroups (ie, myopic sphere for both surgeons, myopic cylinder for both surgeons, hyperopic sphere for both surgeons, and hyperopic cylinder for Surgeon A), as shown in Table 5.

Comparing the total dioptric power between surgeons, there was a small but statistically significant difference of 0.06 D for myopia and 0.08 D for hyperopia, indicating that Surgeon A tended to find a larger difference in refraction (compared with the optometrist refraction) than Surgeon B for myopia and hyperopia. Comparing sphere between surgeons, there was a small but statistically significant difference of 0.14 D for myopia and 0.04 D for hyperopia. Surgeon B tended to push more plus than Surgeon A in myopia but not in hyperopia. Comparing cylinder between surgeons, there was no statistically significant difference. Finally, comparing the SD for total dioptric power between surgeons, there was a statistically significant difference in total dioptric power, sphere, and cylinder in the myopia and hyperopia groups, with a higher SD for Surgeon A in all cases.

The mean surgical outcome accuracy of the SE refraction after LASIK was −0.12 ± 0.42 D for Surgeon A and −0.18 ± 0.43 D for Surgeon B in the myopia group and +0.09 ± 0.57 D and 0.00 ± 0.57 D, respectively, in the hyperopia group. The mean follow-up was 12 months in each group.

Figure 2 shows a Bland-Altman plot of the mean SE refraction against the difference in SE refraction in all eyes. The plot shows that variability did not increase at the extremes of refractive error and was consistent between myopic eyes and hyperopic eyes.

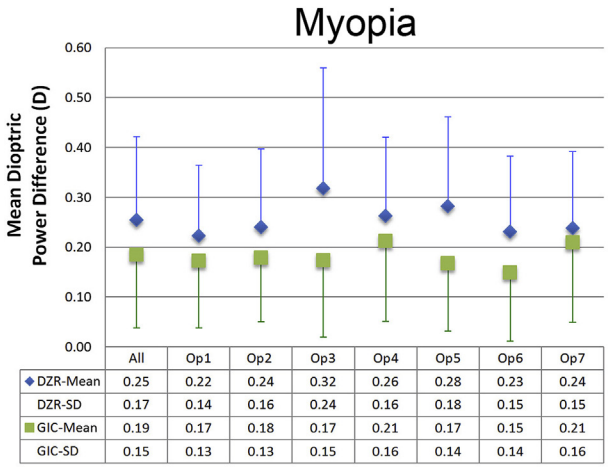
Figure 3 shows a scatterplot of the time between visits against the difference in SE refraction between the optometrists and the surgeons in the myopic group. The refractive stability was independent of the time between visits, although the 2 outliers with the greatest difference (myopic shift) between visits were refractions performed 60 days apart. Figure 4 shows a scatterplot of the time between visits against the difference in SE refraction between the optometrists and the surgeons in the hyperopia group. This also showed that the refractive stability was independent of the time between visits.

DISCUSSION

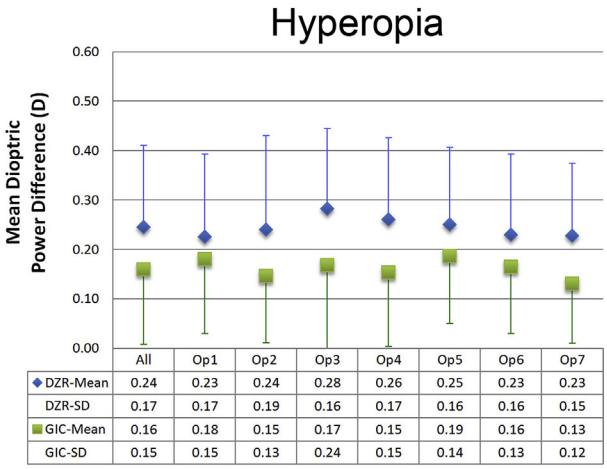
The present study found a consistently high interobserver reproducibility of manifest refraction between 7 optometrists and each of 2 surgeons in routine clinical practice where all practitioners follow a specific and standardized protocol for manifest refraction. Including all optometrists and surgeons, the interobserver median dioptric power difference of 0.19 D (94% ≤0.50 D) compared favorably with the intraobserver repeatability and interobserver reproducibility of manifest refraction that has been reported under controlled conditions.[7,8,10,13]

Raasch et al.[7] report a median difference in total dioptric power of 0.20 D (95% ≤0.62 D) for a single observer performing repeated manifest refractions in 40 eyes 1 to 14 days apart under controlled conditions, which is very similar to the results in the present study (95% ≤0.52 D). Rosenfield and Chiu[8] report the 95% LoA for SE of ±0.29 D for a single observer performing 5 manifest refractions in 12 eyes at the same time of day on different days over a 2-week period under controlled conditions. The equivalent statistic for SE refraction in the present study was only ±0.45 D, even though our study had multiple observers with the refractions occurring at larger time intervals in a routine clinical practice for a far greater number of eyes.

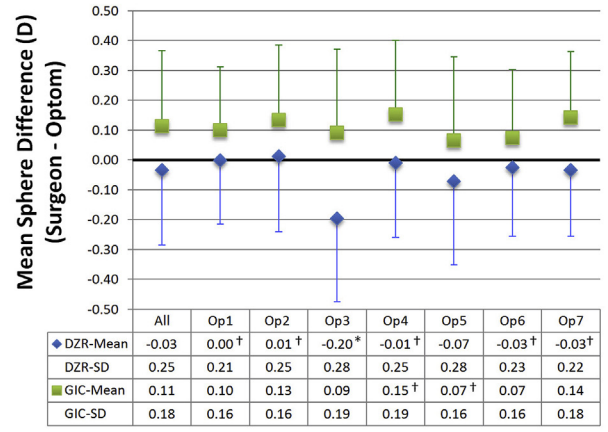
Bullimore et al.[13] measured interobserver reproducibility of refraction. Two examiners performed a manifest refraction back-to-back in 86 subjects (172 eyes) and found a mean difference in SE refraction of



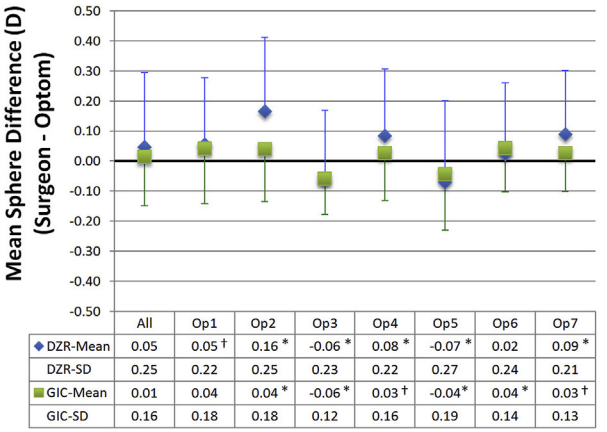
	All	Op1	Op2	Op3	Op4	Op5	Op6	Op7
DZR-Mean	0.25	0.22	0.24	0.32	0.26	0.28	0.23	0.24
DZR-SD	0.17	0.14	0.16	0.24	0.16	0.18	0.15	0.15
GIC-Mean	0.19	0.17	0.18	0.17	0.21	0.17	0.15	0.21
GIC-SD	0.15	0.13	0.13	0.15	0.16	0.14	0.14	0.16



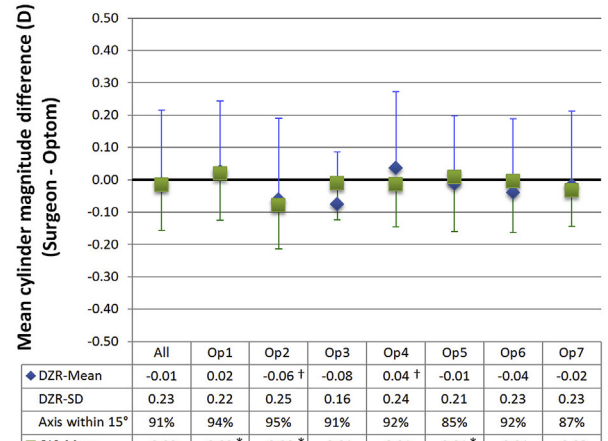
	All	Op1	Op2	Op3	Op4	Op5	Op6	Op7
DZR-Mean	0.24	0.23	0.24	0.28	0.26	0.25	0.23	0.23
DZR-SD	0.17	0.17	0.19	0.16	0.17	0.16	0.16	0.15
GIC-Mean	0.16	0.18	0.15	0.17	0.15	0.19	0.16	0.13
GIC-SD	0.15	0.15	0.13	0.24	0.15	0.14	0.13	0.12



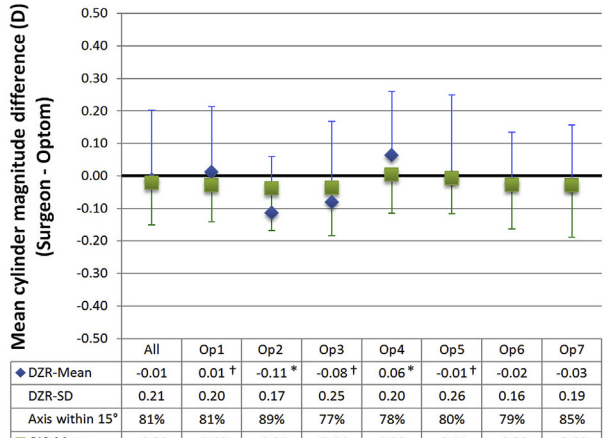
	All	Op1	Op2	Op3	Op4	Op5	Op6	Op7
DZR-Mean	-0.03	0.00 †	0.01 †	-0.20 *	-0.01 †	-0.07	-0.03 †	-0.03 †
DZR-SD	0.25	0.21	0.25	0.28	0.25	0.28	0.23	0.22
GIC-Mean	0.11	0.10	0.13	0.09	0.15 †	0.07 †	0.07	0.14
GIC-SD	0.18	0.16	0.16	0.19	0.19	0.16	0.16	0.18



	All	Op1	Op2	Op3	Op4	Op5	Op6	Op7
DZR-Mean	0.05	0.05 †	0.16 *	-0.06 *	0.08 *	-0.07 *	0.02	0.09 *
DZR-SD	0.25	0.22	0.25	0.23	0.22	0.27	0.24	0.21
GIC-Mean	0.01	0.04	0.04 *	-0.06 *	0.03 †	-0.04 *	0.04 *	0.03 †
GIC-SD	0.16	0.18	0.18	0.12	0.16	0.19	0.14	0.13



	All	Op1	Op2	Op3	Op4	Op5	Op6	Op7
DZR-Mean	-0.01	0.02	-0.06 †	-0.08	0.04 †	-0.01	-0.04	-0.02
DZR-SD	0.23	0.22	0.25	0.16	0.24	0.21	0.23	0.23
Axis within 15°	91%	94%	95%	91%	92%	85%	92%	87%
GIC-Mean	-0.02	0.02 *	-0.08 *	-0.01	-0.01	0.01 *	-0.01	-0.03
GIC-SD	0.14	0.14	0.13	0.11	0.13	0.17	0.16	0.11
Axis within 15°	97%	97%	96%	97%	97%	99%	100%	97%



	All	Op1	Op2	Op3	Op4	Op5	Op6	Op7
DZR-Mean	-0.01	0.01 †	-0.11 *	-0.08 †	0.06 *	-0.01 †	-0.02	-0.03
DZR-SD	0.21	0.20	0.17	0.25	0.20	0.26	0.16	0.19
Axis within 15°	81%	81%	89%	77%	78%	80%	79%	85%
GIC-Mean	-0.02	-0.03	-0.04	-0.04	0.00	-0.01	-0.03	-0.03
GIC-SD	0.13	0.11	0.13	0.15	0.12	0.11	0.14	0.16
Axis within 15°	93%	89%	100%	93%	92%	98%	84%	95%

A *B*

Figure 1. Mean and SD for the difference in total dioptric power, sphere, and cylinder magnitude for each optometrist–surgeon combination in the myopic group and hyperopic group. Variability in axis is also included. A dagger (†) indicates a statistically significant difference with 1 other optometrist; an asterisk (*) indicates a statistically significant difference with 2 or more other optometrists (DZR = Surgeon *A*; GIC = Surgeon *B*; Optom = optometrist).

Table 3. Percentage accuracy of dioptric powers.

	All	Opt 1	Opt 2	Opt 3	Opt 4	Opt 5	Opt 6	Opt 7
Myopia								
Surgeon A								
Eyes (n)	487	76	87	39	110	76	44	55
Within ±0.25 D (%)	52	64	54	36	51	47	50	51
Within ±0.50 D (%)	91	96	90	77	92	91	95	89
Surgeon B								
Eyes (n)	467	71	60	41	102	73	44	76
Within ±0.25 D (%)	71	65	77	78	65	74	80	67
Within ±0.50 D (%)	96	100	100	95	94	96	98	93
Hyperopia								
Surgeon A								
Eyes (n)	492	89	70	40	112	89	33	59
Within ±0.25 D (%)	56	62	60	45	55	53	55	59
Within ±0.50 D (%)	91	91	91	88	89	91	94	92
Surgeon B								
Eyes (n)	476	67	63	47	107	64	53	75
Within ±0.25 D (%)	76	70	81	81	73	61	83	84
Within ±0.50 D (%)	97	94	97	91	97	98	98	99

Opt = optometrist

−0.12 D with a 95% confidence interval (CI) (±2 SD) of 1.55 D (range −0.90 to +0.65 D). In the present study, the mean difference was of smaller magnitude (+0.03 D) with a much narrower 95% CI (0.89 D) (range −0.42 to +0.47 D) despite the larger time interval between refractions, retrospective analysis of refraction in a routine clinical practice (the clinicians were not aware that their measurements would be used to study reproducibility of refraction), and a much greater number of eyes. However, the observers in Bullimore et al. did not have other information, such as spectacle refraction or autorefraction and sphere was masked in their study. MacKenzie[10] used 40 optometrists to perform a manifest refraction for a single patient and reported 95% LoA of ±0.55 D for SE refraction compared with ±0.45 D in the present study.

The range of mean dioptric power difference was between 0.15 D and 0.32 D for all optometrist–surgeon combinations, showing the consistency of the performance of the optometrists and surgeons, which compared with other published studies provides evidence that the use of a standardized protocol for manifest refraction is beneficial to interobserver reproducibility. The consistency across observers is further emphasized by the 1-way ANOVA results, which found no statistically significant difference in mean total dioptric power values in the myopia group and hyperopia group between optometrists compared with the results of either surgeon. In most other subgroup analyses, the ANOVA showed a statistically significant difference. However, these differences were clinically insignificant and were the result of 1 or 2 of the 7 optometrists as described below.

For sphere in myopic eyes compared with Surgeon A, optometrist 3 tended to measure 0.20 D less myopia; however, the maximum difference between optometrists was only 0.21 D. This difference was examined and may have been biased by several factors resulting from nonrandomization of subjects. This group (optometrist 3 to Surgeon A) included the only eye in which a difference of −1.00 D occurred; this eye was highly myopic in which the 2 manifest refractions were performed 2 months apart. This group also had the highest proportion of eyes in which the manifest refractions were performed more than 1 month apart (33%) and had the second fewest number of eyes. For sphere in myopic eyes compared with Surgeon B, the only statistically significant difference was between optometrists 4 and 5, and the difference was only 0.08 D.

For sphere in hyperopic eyes compared with Surgeon A, optometrist 3 and optometrist 5 were the only two to have a negative mean (ie, the optometrist measured more hyperopia than the surgeon). However, the largest difference was only 0.23 D between optometrist 2 and optometrist 5. Similarly, for sphere in hyperopic eyes compared with Surgeon B, optometrist 3 and optometrist 5 were again the only 2 to have a negative mean (optometrist measured more hyperopia than surgeon), although the largest difference was only 0.10 D between optometrist 2 and optometrist 3.

Table 4. Mean difference and 95% LOA for M, J0, J45 and cylinder vector difference.

Group/Surgeon	Mean (Limits of Agreement)							
	All	Opt 1	Opt 2	Opt 3	Opt 4	Opt 5	Opt 6	Opt 7
Myopia								
Surgeon A								
Eyes (n)	487	76	87	39	110	76	44	55
M (D)	−0.04 (±0.37)	+0.01 (±0.42)	−0.02 (±0.46)	−0.23 (±0.56)	+0.01 (±0.49)	−0.08 (±0.53)	−0.05 (±0.45)	−0.04 (±0.46)
J0 (D)	−0.02 (±0.24)	−0.02 (±0.19)	0.00 (±0.28)	−0.01 (±0.19)	−0.03 (±0.26)	−0.04 (±0.24)	+0.01 (±0.21)	−0.01 (±0.23)
J45 (D)	−0.01 (±0.21)	−0.01 (±0.22)	−0.01 (±0.16)	−0.03 (±0.23)	−0.02 (±0.22)	+0.02 (±0.25)	−0.01 (±0.22)	−0.02 (±0.19)
Cylinder (D)	0.27 (±0.37)	0.26 (±0.31)	0.25 (±0.41)	0.25 (±0.35)	0.29 (±0.38)	0.29 (±0.41)	0.24 (±0.36)	0.25 (±0.35)
Surgeon B								
Eyes (n)	467	71	60	41	102	73	44	76
M (D)	+0.11 (±0.36)	+0.11 (±0.31)	+0.09 (±0.32)	+0.09 (±0.40)	+0.14 (±0.40)	+0.07 (±0.32)	+0.07 (±0.32)	+0.13 (±0.40)
J0 (D)	−0.01 (±0.14)	−0.02 (±0.14)	−0.01 (±0.16)	−0.01 (±0.14)	−0.01 (±0.14)	0.00 (±0.15)	−0.02 (±0.14)	+0.01 (±0.15)
J45 (D)	0.00 (±0.15)	0.00 (±0.15)	−0.01 (±0.16)	+0.01 (±0.08)	0.00 (±0.14)	0.00 (±0.20)	0.00 (±0.12)	−0.02 (±0.16)
Cylinder (D)	0.14 (±0.32)	0.14 (±0.29)	0.16 (±0.32)	0.10 (±0.25)	0.14 (±0.29)	0.15 (±0.38)	0.12 (±0.30)	0.14 (±0.35)
Hyperopia								
Surgeon A								
Eyes (n)	492	89	70	40	112	89	33	59
M (D)	+0.04 (±0.47)	+0.06 (±0.46)	+0.11 (±0.47)	−0.10 (±0.43)	+0.11 (±0.45)	−0.07 (±0.43)	+0.01 (±0.49)	+0.07 (±0.43)
J0 (D)	−0.02 (±0.23)	−0.02 (±0.22)	−0.02 (±0.21)	−0.02 (±0.33)	−0.02 (±0.23)	−0.05 (±0.26)	−0.03 (±0.15)	−0.02 (±0.21)
J45 (D)	−0.01 (±0.22)	−0.01 (±0.19)	0.00 (±0.24)	−0.01 (±0.28)	−0.02 (±0.24)	−0.01 (±0.23)	0.00 (±0.21)	−0.01 (±0.19)
Cylinder (D)	0.27 (±0.40)	0.24 (±0.33)	0.25 (±0.41)	0.32 (±0.58)	0.28 (±0.36)	0.28 (±0.46)	0.22 (±0.29)	0.24 (±0.31)
Surgeon B								
Eyes (n)	476	67	63	47	107	64	53	75
M (D)	0.00 (±0.35)	+0.03 (±0.38)	+0.02 (±0.35)	−0.08 (±0.30)	+0.03 (±0.33)	−0.05 (±0.39)	+0.03 (±0.34)	+0.01 (±0.28)
J0 (D)	−0.01 (±0.19)	+0.01 (±0.12)	−0.01 (±0.13)	−0.03 (±0.40)	−0.01 (±0.15)	0.00 (±0.13)	−0.03 (±0.17)	−0.01 (±0.15)
J45 (D)	0.00 (±0.17)	−0.01 (±0.22)	0.00 (±0.09)	−0.01 (±0.21)	−0.01 (±0.15)	+0.01 (±0.15)	0.00 (±0.15)	−0.01 (±0.14)
Cylinder (D)	0.15 (±0.41)	0.16 (±0.39)	0.11 (±0.24)	0.20 (±0.83)	0.16 (±0.40)	0.14 (±0.31)	0.15 (±0.36)	0.15 (±0.28)

J0 = *x*-coordinate of the astigmatism vector; J45 = *y*-coordinate of the astigmatism; M = spherical equivalent refraction; LoA = limits of agreement; Opt = optometrist

For cylinder magnitude in myopic eyes compared with Surgeon A, the only statistically significant difference was between optometrists 2 and 4; however, the difference was only 0.10 D. For cylinder magnitude in myopic eyes compared with Surgeon B, the only statistically significant differences were between optometrists 1, 2, and 5; however, again the maximum difference was only 0.10 D. For cylinder magnitude in hyperopic eyes compared with Surgeon A, there were statistically significant differences for 7 optometrist pairs involving 5 optometrists; however, the maximum difference was only 0.17 D.

Therefore, even though ANOVA showed statistically significant differences in a few cases, these were clinically negligible because the optometrist-pair difference was between 0.08 D and 0.23 D—less than the minimum measurement interval of refraction itself. The magnitude of these differences is comparable to previous intraobserver repeatability and interobserver reproducibility studies under controlled conditions, which would normally be expected to have even higher reproducibility.

There was a greater difference between the 7 optometrists' manifest refractions and Surgeon A than Surgeon B; the mean total dioptric power difference was 0.25 D for Surgeon A and 0.19 D for Surgeon B in the myopic group and 0.24 D and 0.16 D, respectively, in the hyperopic group. Although Surgeon B tended to measure less myopic sphere than the optometrists and Surgeon A (mean difference 0.14 D), Surgeon A was more likely to change the refraction as shown by the fact that the SD of the difference between the optometrist and surgeon refractions was higher for both sphere and cylinder. However, although these differences were statistically significant, there was little clinical significance as shown by the statistically identical accuracy of the refractive outcome after LASIK between surgeons. Also, neither Surgeon A nor Surgeon B was identified as a statistically significant variable in our multivariate nomogram, which had been updated in April 2011.[A] It is likely therefore that the biological noise in the healing process of the procedure overshadows these small statistically detected differences.

An additional dimension to this study was the Bland-Altman chart to show graphically whether the magnitude of refractive error had an effect on the reproducibility of refraction. The presence of very few outliers in such a large sample and no visible

Table 5. Results of the one-way ANOVA and Tukey post hoc test performed across all 7 optometrists compared with 2 surgeons for total dioptric power, sphere, and cylinder magnitude. The table includes all optometrist pairs, out of the 21 possible pairs, for which the Tukey post hoc test found a statistically significant difference.

Surgeon	Ametropia	Refraction	*P* Value		Difference (D)
			ANOVA	Tukey Post Hoc Test	
Surgeon A	Myopia (487 eyes)	Total dioptric power	.050	(0 of 21)	
		Sphere	.001*	(5 of 21)	
				Opt 3–Opt 1 (.002)	0.20
				Opt 3–Opt 2 (<.001)	0.21
				Opt 3–Opt 4 (.001)	0.19
				Opt 3–Opt 6 (.031)	0.17
				Opt 3–Opt 7 (.032)	0.17
		Cylinder magnitude	.021*	(1 of 21)	
				Opt 2–Opt 4 (.042)	0.10
Surgeon B	Myopia (467 eyes)	Total dioptric power	.109	(0 of 21)	
		Sphere	.012*	(1 of 21)	
				Opt 4–Opt 5 (.022)	0.08
		Cylinder magnitude	.003*	(3 of 21)	
				Opt 1–Opt 2 (.001)	0.10
				Opt 1–Opt 5 (.008)	0.01
				Opt 2–Opt 5 (.007)	0.09
Surgeon A	Hyperopia (492 eyes)	Total dioptric power	.507	(0 of 21)	
		Sphere	<.001*	(7 of 21)	
				Opt 1–Opt 5 (.008)	0.12
				Opt 2–Opt 3 (<.001)	0.22
				Opt 2–Opt 5 (<.001)	0.23
				Opt 3–Opt 4 (.015)	0.14
				Opt 3–Opt 7 (.030)	0.15
				Opt 4–Opt 5 (.001)	0.15
				Opt 5–Opt 7 (.001)	0.06
		Cylinder magnitude	.001*	(4 of 21)	
				Opt 2–Opt 1 (.003)	0.12
				Opt 2–Opt 4 (<.001)	0.17
				Opt 2–Opt 5 (.020)	0.10
				Opt 3–Opt 4 (.004)	0.14
Surgeon B	Hyperopia (476 eyes)	Total dioptric power	.337	(0 of 21)	
		Sphere	.001*	(6 of 21)	
				Opt 3–Opt 2 (.012)	0.10
				Opt 3–Opt 4 (.023)	0.09
				Opt 3–Opt 6 (.014)	0.10
				Opt 3–Opt 7 (.037)	0.09
				Opt 5–Opt 2 (.032)	0.08
				Opt 5–Opt 6 (.036)	0.08
		Cylinder magnitude	.319	(0 of 21)	

ANOVA = analysis of variance; Opt = optometrist
*Statistically significant

trends shows that the reproducibility was independent of the magnitude of the refractive error.

One difference between conditions at the time of the first refraction and the second refraction was that optometrists at the first refraction had the current spectacle refraction (if available) and aberrometric autorefraction before performing the manifest refraction, while surgeons had this information in addition to the optometrists' manifest and cycloplegic refractions before starting their manifest refraction. Therefore, the results in this study cannot be strictly interpreted in comparison with a rigorous double-masked controlled study. The surgeons also had the benefit of seeing the optometrists' cycloplegic

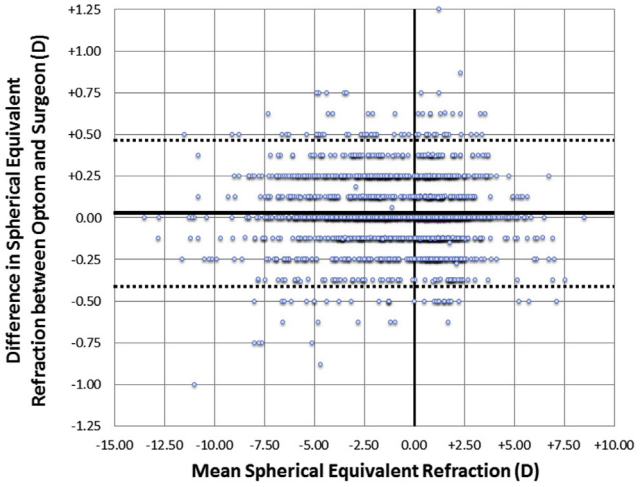


Figure 2. Bland-Altman plot showing the variation of differences in SE. The mean SE and ±1.96 SD are displayed as lines on the chart. No trend is visible at the extremes of refraction (Optom = optometrist).

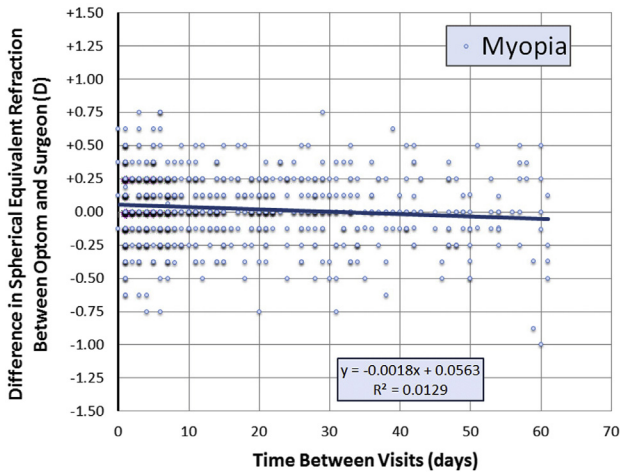


Figure 3. Time between visits against the difference in SE refraction for all myopic eyes. There was similar stability in SE refraction for the range of 1 to 62 days between visits (Optom = optometrist).

refraction, whereas the optometrists performed their manifest refraction without knowledge of the cycloplegic refraction values. The presence of a clinically significant difference between these 2 refractions, for example if the cycloplegic refraction were more hyperopic than the manifest refraction, might have caused the surgeon to push more plus, resulting in a more positive result than that of the optometrist, which was observed for 10 out of 14 surgeon/optometrist combinations for measurement of hyperopic sphere, with the same 2 optometrists measuring slightly more plus than both surgeons. However, the difference was no greater than 0.16 D, which suggests that the optometrists were successful in

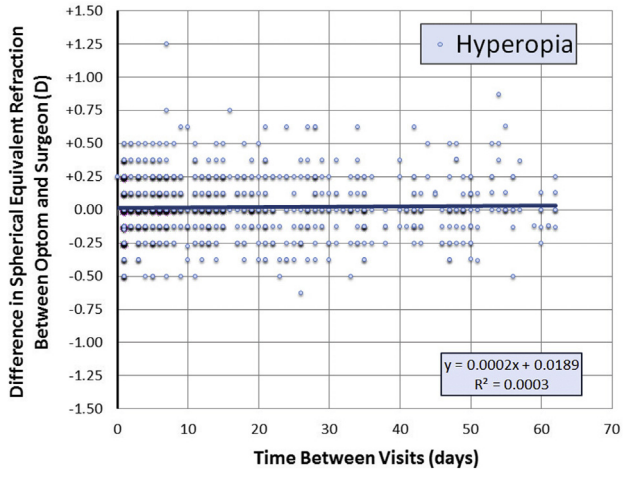


Figure 4. Time between visits against the difference in SE refraction for all hyperopic eyes. There was similar stability in spherical SE for the range of 1 to 62 days between visits (Optom = optometrist).

"pushing plus" during the first manifest refraction. Nevertheless, surgeon bias is unlikely because the surgeons were refracting the patients with a view to performing laser refractive surgery, the motivation was to achieve the most accurate manifest refraction possible rather than to match the optometrist manifest refraction.

Another weakness of the present study was that the time between the 2 manifest refractions was not fixed for all patients, ranging between 1 day and 2 months, and the 2 manifest refractions were not performed at the same time of day. This introduces some potential pseudo errors in manifest refraction due to actual natural and physiologic fluctuation in refraction over time and/or diurnally. The ideal study design would have been for the 2 manifest refractions to be performed on the same day without reference to any other refraction and with equal intervals between the refractions. However, the aim of this study was to evaluate the interobserver reproducibility given and despite the above variations, to assess the potential accuracy of refraction for planning refractive surgery in our clinical refractive surgery practice.

In conclusion, the present study found that stringent training with a standardized manifest refraction protocol resulted in consistent interobserver reproducibility independent of the magnitude of refractive error. We found that we could achieve interobserver reproducibility in our everyday clinical practice that compared favorably with the intraobserver repeatability or interobserver reproducibility achieved in fully controlled studies.

WHAT WAS KNOWN

- Intraobserver repeatability and interobserver reproducibility of manifest refraction under controlled conditions for small populations have been reported to be in the region of 0.25 D.

WHAT THIS PAPER ADDS

- Interobserver repeatability of manifest refraction between multiple practitioners in routine clinical practice can be made to be similar to that of the intraobserver repeatability of controlled studies by using a standardized refraction protocol.

REFERENCES

1. Mrochen M, Hafezi F, Iseli HP, Löffler J, Seiler T. Verbesserung der refraktiven Ergebnisse durch Nomogramme [Nomograms for the improvement of refractive outcomes]. Ophthalmologe 2006; 103:331–338
2. Yang SH, Van Gelder RN, Pepose JS. Neural network computer program to determine photorefractive keratectomy nomograms. J Cataract Refract Surg 1998; 24:917–924
3. Huang D, Stulting RD, Carr JD, Thompson KP, Waring GO III. Multiple regression and vector analyses of laser in situ keratomileusis for myopia and astigmatism. J Refract Surg 1999; 15:538–549
4. Ditzen K, Handzel A, Pieger S. Laser in situ keratomileusis nomogram development. J Refract Surg 1999; 15:S197–S201
5. Biebesheimer JB, Kang TS, Huang CY, Yu F, Hamilton DR. Development of an advanced nomogram for myopic astigmatic wavefront-guided laser in situ keratomileusis (LASIK). Ophthalmic Surg Lasers Imaging 2011; 42:241–247
6. Liyanage SE, Allan BD. Multiple regression analysis in myopic wavefront laser in situ keratomileusis nomogram development. J Cataract Refract Surg 2012; 38:1232–1239
7. Raasch TW, Schechtman KB, Davis LJ, Zadnik K; and the CLEK Study Group. Repeatability of subjective refraction in myopic and keratoconic subjects: results of vector analysis. Ophthalmic Physiol Opt 2001; 21:376–383
8. Rosenfield M, Chiu NN. Repeatability of subjective and objective refraction. Optom Vis Sci 1995; 72:577–579. Available at: http://journals.lww.com/optvissci/Abstract/1995/08000/Repeatability_of_Subjective_and_Objective.7.aspx. Accessed October 18, 2013
9. Shah R, Edgar DF, Rabbetts R, Harle DE, Evans BJW. Standardized patient methodology to assess refractive error reproducibility. Optom Vis Sci 2009; 86:517–528. Available at: http://journals.lww.com/optvissci/Fulltext/2009/05000/Standardized_Patient_Methodology_to_Assess.14.aspx. Accessed October 18, 2013
10. MacKenzie GE. Reproducibility of sphero-cylindrical prescriptions. Ophthalmic Physiol Opt 2008; 28:143–150
11. Reinstein DZ, Archer TJ, Couch D. Accuracy of the WASCA aberrometer refraction compared to manifest refraction in myopia. J Refract Surg 2006; 22:268–274
12. Thibos LN, Horner D. Power vector analysis of the optical outcome of refractive surgery. J Cataract Refract Surg 2001; 27:80–85
13. Bullimore MA, Fusaro RE, Adams CW. The repeatability of automated and clinician refraction. Optom Vis Sci 1998; 75:617–622. Available at: http://journals.lww.com/optvissci/Abstract/1998/08000/The_Repeatability_of_Automated_and_Clinician.28.aspx. Accessed October 18, 2013

OTHER CITED MATERIAL

A. Dan Z. Reinstein, MD, personal communication, May 9, 2011

First author:
Dan Z. Reinstein, MD, MA(Cantab), FRCSC, DABO, FRCOphth, FEBO

London Vision Clinic, London, United Kingdom