Reliability of Noncontact Pachymetry after Laser In Situ Keratomileusis

Miguel J. Maldonado,1,2 Alberto López-Miguel,1 Juan C. Nieto,1 Juan Cano-Parra,3 Beöña Calvo,4,5 and Jorge L. Alió6

PURPOSE. To assess the repeatability and reproducibility of central corneal thickness (CCT) measurements obtained by combined scanning-slit/Placido-disc topography (Orbscan II; Bausch & Lomb, Rochester, NY) after laser in situ keratomileusis (LASIK) and to compare the results with another noncontact method, specular microscopy (SP-2000P; Topcon, Tokyo, Japan).

METHODS. To analyze intrasession repeatability, one examiner measured 22 postmyopic LASIK eyes 10 times successively in the shortest time possible, using both devices randomly. To study intersession reproducibility, the same operator obtained measurements from another 50 eyes with stable refraction in two consecutive visits at the same time of the day between 6 and 9 months after myopic LASIK. Any association between residual stromal bed thickness and measurement variability was recorded and evaluated.

RESULTS. For intrasession repeatability, Orbscan II and Topcon SP-2000P CCT measurements showed a repeatability of 20.2 (4.3%) and 12.8 (2.6%) μm, respectively. Both devices yielded excellent intraocular correlation coefficients (ICC; 0.98 and 0.99, respectively). For intersession reproducibility, no difference in CCT measurements was found. The coefficient of intersession reproducibility was 6.7% (29.5 μm) for Orbscan II and 4.3% (19.5 μm) for Topcon SP-2000P CCT measurements. The ICCs (0.95 and 0.96, respectively) indicated good intersession reliability. Repeatability and reproducibility with both devices were unrelated to stromal bed thickness.

CONCLUSIONS. Both noncontact pachymeters provide repeatable CCT measurements in transparent postmyopic LASIK corneas after the early postoperative period. Intersession variations in CCT of more than 29 μm with the Orbscan II and 19 μm with the Topcon SP-2000P may reflect true corneal change. These estimates should help investigators and clinicians differentiate actual CCT modifications from measurement variability.

SST pachymetry has the advantage over ultrasound pachymetry of not requiring corneal anesthesia and of using a contact probe that in turn reduces iatrogenic epithelial alterations, the risk of microbial contamination, and flap dislocation that can occur, even after minor trauma.26 However, some factors may cause errors in SST measurements, making the reliability of this system questionable.1,20,27 It is therefore crucial to assess the random errors in SST—namely, repeatability and reproducibility—to determine its precision and validity in patients who have undergone LASIK.28,29 In several studies, investigators have evaluated the CCT reliability of the SST alone30 and combined with Placido-disc videokeratography31–32 (Orbscan I and II, respectively; Bausch & Lomb, Rochester, NY) in normal corneas. In addition, the accuracy but not the precision of Orbscan II pachymetry in contrast to ultrasound has been studied in post-LASIK corneas.33 Therefore, there is an absence of information regarding the precision of Orbscan II pachymetry in post-LASIK eyes.

Noncontact specular microscopy (NCSM) is also widely used in clinical practice to measure corneal endothelial cell density and CCT.34 In normal corneas, the intraobserver reproducibility of NCSM has been reported,35 and its accuracy with respect to ultrasound pachymetry has been established.36 However, there are no consistent definitions of reproducibility and repeatability, and to the best of our knowledge, no study has been undertaken to evaluate them specifically in post-LASIK corneas.

---

From the 1Department of Ophthalmology, University Clinic, University of Navarra, Pamplona, Spain; 2IOBA (Institute of Applied Ophthalmobiology), University of Valladolid, Valladolid, Spain; the 3Municipal Hospital Badalona, Barcelona, Spain; the 4Group of Structural Mechanics and Materials Modeling (GEMM), Aragon Institute of Engineering Research (I3A), University of Zaragoza, Zaragoza, Spain; the 5Networking Center on Bioengineering, Biomaterials and Nanomedicine (CIBER-BBN), Aragon Institute of Health (ICS), Zaragoza, Spain; and the 6Vissum-Instituto Oftalmológico de Alicante, Universidad Miguel Hernández, Alicante, Spain.

Supported in part by RETICS RD07/0062 (Oftalmología), by Plan Nacional 1+D+I, Proyecto Coordinado FIS2005-05020-C03-03, Instituto de Salud Carlos III, Madrid, Spain.

Submitted for publication January 14, 2009; revised February 27 and March 18, 2009; accepted July 13, 2009.

Disclosure: M.J. Maldonado, None; A. López-Miguel, None; J.C. Nieto, None; J. Cano-Parra, None; B. Calvo, None; J.L. Alió, None;

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Corresponding author: Miguel J. Maldonado, Department of Ophthalmology, Clinica Universitaria, Av. Pio XII, 36, 31080 Pamplona, Spain; mjmalдонадо@unav.es.

One purpose of this study was to quantify for the first time both the reproducibility and repeatability of the Orbscan II topographer and Topcon NCSM (SP-2000P specular pachometer; Topcon Corp., Tokyo, Japan) in eyes treated with LASIK using definitions of reproducibility and repeatability based on those of the International Standards Organization, as recommended by Bland and Altman. It is also worth assessing the performance of the equipment when the residual posterior corneal stromal bed is of borderline thickness, because this subset of eyes may be more prone to structural instability. Thus, another study objective was analysis of the precision of these two systems in post-LASIK corneas regarding the stromal bed thickness.

**Materials and Methods**

All procedures were performed in accordance with the Declaration of Helsinki. All candidates received detailed information about the nature of the investigation, and all provided informed consent. The study was approved by The University of Navarra Clinic Review Committee. We selected consecutive patients who had refractive stability on two successive visits at least 2 months apart.

The patients were recruited over a 2-year period. Exclusion criteria consisted of suspected keratocoma, progressive myopia, or astigmatism (patients with a change of 0.25 D and/or ≥15° in cycloplegic refraction cylinder orientation were excluded), preoperative or post-LASIK hyperopia, mixed astigmatism, active ocular disease, connective tissue disorders, and pregnancy. All eyes underwent a complete ophthalmic examination before and after LASIK, including manifest and cycloplegic refractions, Orbscan II pachymetry, tangential videokeratography, noncontact specular pachymetry, slit lamp microscopy, applanation tonometry, and indirect ophthalmoscopy.

One experienced technician was assigned to each instrument to minimize variations in the results. The examiners were masked to the refraction and the postoperative time interval at each examination. Central pachymetry also was performed by Topcon NCSM. Noncontact specular pachymetry was performed immediately before or after Orbscan II pachymetry; the order of the modalities was random.

During LASIK follow-up, objective measurement of corneal transparency was performed with digital imaging analysis of anterior slit lamp photographs, as previously reported. We also assessed the central stromal bed thickness using optical coherence tomography. The measurements were confirmed to be within 10 μm during LASIK enhancement with a portable ultrasound pachometer (Corneo-Gage Plus II; Sonogage Inc, Cleveland, OH). One eye of each patient was selected randomly for this study. No patient was lost to follow-up.

**Study Design**

Our definitions of reproducibility and repeatability were based on those adopted by the British Standards Institution and the International Standards Organization.

**Intrasession Repeatability.** Under repeatability conditions, independent test results were obtained using the same method, on the same subject, with the same operator, on the same equipment with the shortest time possible between successive sets of readings. We investigated the repeatability by performing 10 examinations of 22 eyes after ensuring proper focusing of the entire cornea.

**Intersession Reproducibility.** Under reproducibility conditions, sets of readings were obtained by using the same method, but with one variation in the experimental set up. In this part of the study, we investigated intersession reproducibility in 50 eyes by performing examinations 6 and 9 months after surgery by the same experienced operator. Examinations were performed during a specific period (11 AM to 1 PM) to minimize the effect of diurnal variations in corneal thickness.

**Surgical Procedures**

All procedures were performed by the same surgeon (MJM). The details of the primary LASIK procedures have been reported. The preoperative CCT had to be adequate to ensure a residual stromal bed of at least 250 μm over 55% of the total preoperative CCT after creating a superiorly hinged corneal flap (160-μm head; Hansatome microkeratome; Chiron Vision, Claremont, CA). Ablation of the appropriate cut depth was performed with an excimer laser (Technolas Keracor 217z; Bausch & Lomb-Chiron Technolabs GmbH, Munich, Germany).

**CCT Evaluation**

The same examiner calibrated the Orbscan II unit before each session. The patient was instructed to fixate on a flickering red light. Before proceeding, the patient was instructed to blink to obtain a homogeneous tear film layer. The instrument then was aligned and the cornea was scanned by a slit beam. The Orbscan II statistical analysis device (software ver. 3.12) calculated the elevation of the anterior and posterior corneal surfaces; thus, the pachymetry is the result of subtracting the elevation map of both corneal surfaces. Although Orbscan II displays pachymetry values at several locations, only the central pachymetry measurement was analyzed because it enabled comparison with Topcon NCSM, which provides CCT values. The default acoustic equivalent setting recommended by the manufacturer was used throughout the study.

For Topcon NCSM pachymetry, the patient was positioned with the chin in a cup and forehead against a headband. The CCT was measured while the subject focused on a fixation light in the instrument. After proper positioning of the alignment dot, circle, and bar on the screen, pachymetry was performed. The automated image capture, low-intensity mode of the specular microscope was used as described previously.

**Statistical Analyses**

Data from the prospectively completed forms were entered into a database, and statistical calculations were performed (SPSS, ver. 11.0 for Windows; SPSS, Chicago, IL). The mean and SD were calculated for normally distributed data. When data did not correspond to a normal distribution, the median (50th percentile) and interquartile range (IR, values between the 25th and 75th percentiles of the distribution) were used. For all statistical tests, a two-tailed P < 0.05 was considered significant.

**Intrasession Repeatability.** We calculated the within-subject SD (s_w) of the 10 consecutive measurements performed on each eye by obtaining the square root of the value called the residual mean square in the one-way analysis of variance. In addition, we expressed the repeatability as a percentage of the mean measurement for each category of measurement (coefficient of repeatability = 2.77 × s_w/mean measurement). We also calculated the within-subject coefficient of variation (CV_w) and the intrasession reliability of the measurement method (intraclass correlation coefficient; ICC). We also calculated the within-subject coefficient of variation (CV_w) and the intrasession reliability of the measurement method (intraclass correlation coefficient; ICC). The paired t-test was used to establish whether there was a significant systematic bias between measurements. The intersession reliability of the measurement method was calculated with the ICC. We obtained correlation coefficients to explore any association between variables.
Table 1. Intrasection Repeatability

<table>
<thead>
<tr>
<th>Pachymetry Method</th>
<th>Overall Mean (Range)</th>
<th>( s_w ) (95% CI)</th>
<th>Repeatability (Coefficient of R)</th>
<th>( CV_w ) (%) (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Orbscan II SST pachymetry (( \mu m ))</td>
<td>474.9 (380/548)</td>
<td>7.31 (6.59/8.05)</td>
<td>20.2 (4.3%)</td>
<td>1.5 (1.3/1.8)</td>
</tr>
<tr>
<td>Topcon NCSM pachymetry (( \mu m ))</td>
<td>482.8 (409/539)</td>
<td>4.61 (4.15/5.06)</td>
<td>12.8 (2.6%)</td>
<td>1.0 (0.6/1.2)</td>
</tr>
</tbody>
</table>

**RESULTS**

**Intrasection Repeatability**

Repeatability was studied in 22 eyes of 22 patients (7 men, 15 women) with an average age of 32.9 ± 7.6 years (range, 24–49). The mean preoperative myopia was \(-5.1 \pm 2.7 \, \text{D} \) (range, \(-0.5 \) to \(-9.75 \) and median astigmatism was \(-0.5 \, \text{D} \) (IR, 0 to \(-1.25 \)). The median residual myopia was \(-0.5 \, \text{D} \) (IR, \(-0.25 \) to \(-1.0 \)) and astigmatism \(-0.5 \, \text{D} \) (IR, \(-0.5 \) to \(-0.75 \)). The mean stromal bed thickness was 351.5 ± 37.3 \( \mu m \) (range, 278–408). Table 1 shows the overall average CCTs corresponding to the repeated measures, the intrasession \( s_w \) and \( CV_w \), and the repeatability. We found no association between the standard deviation and the mean of the repeated measurements (\( P = 0.5 \) and \( P = 0.7 \) for Orbscan II and Topcon NCSM, respectively), and no significant association between the stromal bed thickness and the SD (\( P = 0.3 \) and \( P = 0.7 \) for Orbscan II and Topcon NCSM, respectively).

**Intersession Reproducibility**

Intersession reproducibility was studied in 50 eyes of 50 patients (23 men, 27 women) with an average age of 34.6 ± 4.5 years (range, 26–48). The median preoperative myopia was \(-4.5 \, \text{D} \) (IR, \(-3.0 \) to \(-6.0 \)) and the median astigmatism was \(-1.0 \, \text{D} \) (IR, 0 to \(-1.5 \)). The median residual myopia was \(-0.5 \, \text{D} \) (IR, 0 to \(-0.75 \)) and astigmatism \(-0.5 \, \text{D} \) (IR, \(-0.5 \) to \(-1.0 \)). The mean stromal bed thickness was 326.4 ± 33.2 \( \mu m \) (range, 264–408 \( \mu m \)). Corneal opacification was almost non-existent and did not differ between the two examinations (median and IR, 0 gray levels [range, 0–3] versus 0 gray levels [range, 0–4], \( P = 0.9 \)). Table 2 shows the overall average CCT values corresponding to the two examinations, the intersession \( s_w \), \( CV_w \), and the reproducibility.

Figure 1 shows the Bland-Altman plots of difference versus mean for the studied Orbscan II and Topcon NCSM pachymetry values that confirmed that the variations were independent of the mean measurement value. Figure 2 shows that the assumption that the distribution of differences is approximately normal is correct. We did not find a significant association between the stromal bed thickness and absolute intersession differences (\( P = 0.06 \); \( r_w = -0.3 \) and \( P = 0.9 \); \( r_w = -0.01 \) for Orbscan II and Topcon NCSM, respectively). Table 3 shows the 95% LoA corresponding to the intersession variability.

No systematic bias was found between measurements in the two consecutive visits with either device (Table 4). In the study of 50 participants, the probability was 80% that the study would detect a difference in the measurements from two consecutive visits if the true difference is that shown in Table 4.

Excellent intersession reliability scores were obtained for CCTs; the ICC and their 95% CIs for intersession reproducibility were 0.95 (0.91–0.97) and 0.96 (0.92–0.98) for Orbscan II and Topcon NCSM, respectively.

**DISCUSSION**

Both routine clinical practice and investigator calculations rely on knowledge of the measurement variability of CCT after LASEK. However, to our knowledge, no investigators have adequately addressed this problem with the devices we tested, which are commonly used in the clinical and research settings. Access to accurate noninvasive pachymetry is essential for reducing the risk of corneal epithelial damage or flap displacement due to probe contact and for avoiding microbial contamination. Despite the importance of knowing the precision (random error) of noncontact CCT after LASEK, the present study is the first appropriate statistical investigation of the reliability of Orbscan II CCT measurements after LASEK. In addition, it was compared with another widely used noncontact pachymetry instrument, the Topcon SP-2000P specular microscope.

Although the Orbscan II is equipped with both Placido-based topography and SST, some limitations of scanning-slit corneal topography have been reported in eyes treated with excimer laser keratorefractive surgery, notably that the accuracy of pachymetry after refractive surgery may be affected when there is loss of corneal transparency. S.9.49 SST measures the corneal thickness by subtracting points on the posterior surface from those on the anterior surface. Because there is substantial underestimation of the corneal thickness by Orbscan pachymetry in the early post-LASIK period, the posterior corneal elevation may not be measured accurately in these eyes. Boscia et al. reported a large disparity between the true curvature of the posterior surface, verified with a radiuscope and the measured curvature of the lens based on measurement of the spherical anterior and posterior surfaces. Furthermore, data from corneas in vivo may be even less accurate due to ocular movement during acquisition. Another shortcoming is that the variability increases from the apex to the periphery in both calibrated test surfaces and human corneas; therefore, the pachymetry may be altered. However, although Orbscan II central pachymetry may not be as accurate as ultrasound, the differences can be reduced greatly by applying...
a proper acoustic factor\(^{33}\) or by the subtraction method.\(^{51}\) On the contrary, repeatability depends solely on the quality of the device itself; consequently, it cannot be modified.

All patients underwent a complete topographic analysis of the cornea, but only CCT measurements were selected for the present study instead of the thinnest point, because the lowest pachymetry variability is in the central cornea\(^{45}\) and because the central pachymetry also was measured with Topcon NCSM, allowing Orbscan II CCTs but not the thinnest points to be compared. The difference between two measurements in the same subject is expected to be less than 2.77 \(s_x\) for 95\% of pairs of observations.\(^{7,46}\) This estimation, termed the repeatability\(^{7,46}\) of the different variables studied, is presented in Table 1 and corroborates the ICC findings; coefficients of repeatability of 4.3\% and 2.6\% indicate very good precision. We did not find a relationship between stromal bed thickness and repeatability scores, implying that whatever the source of the random error, it seems to be unrelated to the residual stroma, and that the precision of both devices’ CCT estimates is independent of stromal bed thickness. The repeatability of the Topcon NCSM appeared to be slightly better than that of the Orbscan II in post-LASIK eyes. The present study yielded excellent intrasession reliability scores (>0.9) for Orbscan II that were comparable to the ICC achieved by the Topcon NCSM. In addition, the Orbscan II CV (Table 1) did not differ markedly from the noncontact specular value, and both indicated that there was outstanding repeatability of the Orbscan II and Topcon NCSM CCT in post-LASIK corneas. Regarding Orbscan II, our reliability outcomes were similar to those of Fam et al.\(^{30}\) and Rainer et al.,\(^{29}\) who have reported ICCs of 0.98 and 0.99, respectively, in normal corneas. Thus, our results appeared to be as good as those obtained in subjects who did not undergo a surgical procedure, although our CV (1.5\%) was slightly less favorable than that (0.67\%) reported by Fam et al.\(^{30}\) Regarding Topcon NCSM, our CVs were more favorable than those described for CCT Topcon NCSM in nonsurgical corneas by Cho and Cheung (7.1\%)\(^{34}\) and Ogbuehi and Almubrad (6.8\%)\(^{35}\) and slightly worse than reported by Uc\c{a}khan et al. (0.2\%).\(^{52}\) Likewise, our ICCs were similar to those previously described for normal corneas—that is, 0.99\(^{34}\) and 0.97.\(^{52}\)

To assess intersession reproducibility, we obtained Orbscan II measurements 6 and 9 months after surgery in eyes with no refractive change. We selected these time points because it is typically during the medium-term follow-up that surgeons consider the stability of the refraction, topography, and pachymetry to decide whether to re-treat\(^{16,17}\) and in the long-term to
collect data for intraocular lens power calculations.10–12 In addition, Orbscan II examinations taken early after LASIK tend to be artificially low for scanning slit pachymetry measurements and high for posterior corneal curvature (PCC) values.13–15 These artifacts may be explained by the acquisition process of the Orbscan.49 The instrument optically measures the anterior and posterior surfaces by triangulating 40 calibrated slit beam surfaces and diffusely reflected camera rays.16 Ray triangulations are calculated using the actual physiologic refraction indices of the eye. The alteration in optical transparency of the anterior cornea corresponding to the edematous flap in the early postoperative period may result in light scatter with a resultant increase in the corneal refractive index.49 As an alternative, Chakrabarti et al.5 suggested that reconstruction algorithms may not apply to post-LASIK corneas with an oblate rather than a prolate shape; another study suggested that alterations in the magnification of the posterior surface due to surgical modification of the anterior curvature and thickness may produce artificial changes in the posterior corneal elevation data.53

When noncontact CCT measurements were compared at 6 and 9 months postoperatively, the mean differences (positive or negative) were not significantly different from 0 (Table 4). It is not surprising that, on average, no difference in CCT measurements was found, indicating that whatever the source of the variability, the interobserver variability is random. Of interest, Orbscan II but not Topcon NCSM showed a weak inverse association between CCT variability and stromal bed thickness that approached but did not reach statistical significance. The Orbscan II and Topcon NCSM coefficient of reproducibility (6.7% and 4.3%, respectively; Table 2) would provide useful guidelines for the variability expected if these noncontact pachymeters are used on two consecutive visits. Therefore, these data indicate that CCT reliability is adequate when measured by Orbscan II and Topcon NCSM in different sessions after LASIK. However, the Orbscan II reproducibility (Table 2) and the 95% LoA (Table 3) were not as satisfactory as Topcon NCSM pachymetry values in the present study and in that of Ogubeji and Almubrad55 on normal corneas (−15 to 17 μm). Nevertheless, the 95% LoA obtained by Rah et al.54 (−48.6 to 54.8 μm) with ultrasound pachymetry after LASIK are wider than our LoAs. Although Marsich and Bullimore55 found a narrower 95% LoA (−10.0 to +16.6 μm) when assessing Orbscan CCT, they analyzed eyes that had not undergone LASIK. Our reproducibility ICCs (0.95 and 0.96) were also slightly lower but in the range of normal corneas (0.98) tested with Topcon NCSM.54

The present study had several limitations. Because Topcon NCSM does not provide the thinnest pachymetry measurement, which in most cases lies in the inferotemporal quadrant,50 we analyzed the reproducibility of CCT. One should expect it would be lower considering the added difficulty in identifying the thinnest point from measurement to measurement. The present study did not address post-LASIK ectasia. However, eyes with ectatic disorders should yield less reproducible CCT measurements as previously shown in keratoconic eyes.52 The present study focused on a typical setting with an experienced operator taking successive examinations. However, we did not study interobserver reproducibility, and the results obtained by novice examiners should be worse than ours. Finally, we evaluated the random error, rather than agreement, when measuring CCT after LASIK for myopia with two noncontact methods under repeatability and reproducibility conditions. It has been reported that Orbscan II CCT measurements clearly underestimate those obtained by ultrasound pachymetry in patients who have undergone LASIK,54 and we found a similar bias with respect to Topcon NCSM. However, such a systematic error could be corrected by adequate instrument calibration by changing the acoustic factor54 or using the subtraction method54 as opposed to random errors, which cannot be corrected and reflect the inherent construction quality of the instrument for the given application.

In conclusion, given that noncontact pachymetry is now used by many ophthalmologists and investigators for the assessment of the LASIK patient in the medium and long-term follow-up, it is important to obtain an estimate of the reproducibility of CCT measurements in these patients to differentiate clinical change from measurement variability. The present study showed that Orbscan II and Topcon NCSM data can be used in patients who undergo LASIK for myopia to obtain a reproducible measure of CCT, and we provide the criteria for a significant change, which would be one exceeding the repeatability/reproducibility (2.77σ) of the estimate. Orbscan II and Topcon NCSM reliably measure CCT, since the random errors were clinically acceptable, although Topcon NCSM provided generally better repeatability and reproducibility scores. The precision of these measures of CCT in the medium-term follow-up of eyes that undergo LASIK for myopia was adequate and independent of the residual bed thickness and, therefore, they are clinically useful in the assessment and longitudinal long-term follow-up of post-LASIK eyes with transparent corneas.

Table 3. 95% LoA

<table>
<thead>
<tr>
<th>Pachymetry Method</th>
<th>Lower LoA (95% CI)</th>
<th>Upper LoA (95% CI)</th>
<th>Width of (95%) LoA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Orbscan II SST pachymetry (μm)</td>
<td>−32.1 (−35.1; 29.1)</td>
<td>26.4 (25.4; 29.4)</td>
<td>58.5</td>
</tr>
<tr>
<td>Topcon NCSM pachymetry (μm)</td>
<td>−19.1 (−21.6; −16.7)</td>
<td>20.3 (17.9; 22.8)</td>
<td>39.4</td>
</tr>
</tbody>
</table>

Corresponding to the two examination sessions.

Table 4. Mean Difference between Intersession Measurements

<table>
<thead>
<tr>
<th>Pachymetry Method</th>
<th>Mean Difference (95% CI)</th>
<th>P*</th>
<th>Minimal Detectable Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Orbscan II SST pachymetry (μm)</td>
<td>−2.84 (−7.08 to −1.40)</td>
<td>0.185</td>
<td>5.91</td>
</tr>
<tr>
<td>Topcon NCSM pachymetry (μm)</td>
<td>0.60 (2.32 to −3.52)</td>
<td>0.679</td>
<td>3.99</td>
</tr>
</tbody>
</table>

Power set at 80%; two-sided 5% significance level. *
Paired t-test.
References


