Repeatability of Measuring Corneal Subbasal Nerve Fiber Length in Individuals With Type 2 Diabetes

Nathan Efron, D.Sc., Katie Edwards, Ph.D., Nicola Roper, B.A.(Oxon.), Nicola Pritchard, B.App.Sc.(Optom), Geoff P. Sampson, Ph.D., Ayda M. Shahidi, B.Sc.(Optom.), Dimitrios Vagenas, Ph.D., Anthony Russell, Ph.D., Jim Graham, Ph.D., Mohammad A. Dabbah, Ph.D., and Rayaz A. Malik, Ph.D.

Purpose: To analyze the repeatability of measuring nerve fiber length (NFL) from images of the human corneal subbasal nerve plexus using semiautomated software.

Methods: Images were captured from the corneas of 50 subjects with type 2 diabetes mellitus who showed varying severity of neuropathy, using the Heidelberg Retina Tomograph 3 with Rostock Corneal Module. Semiautomated nerve analysis software was independently used by two observers to determine NFL from images of the subbasal nerve plexus. This procedure was undertaken on two occasions, 3 days apart.

Results: The intraclass correlation coefficient values were 0.95 (95% confidence intervals: 0.92–0.97) for individual subjects and 0.95 (95% confidence intervals: 0.74–1.00) for observer. Bland-Altman plots of the NFL values indicated a reduced spread of data with lower NFL values. The overall spread of data was less for (a) the observer who was more experienced at analyzing nerve fiber images and (b) the second measurement occasion.

Conclusions: Semiautomated measurement of NFL in the subbasal nerve fiber layer is highly repeatable. Repeatability can be enhanced by using more experienced observers. It may be possible to markedly improve repeatability when measuring this anatomic structure using fully automated image analysis software.

Key Words: Repeatability—Corneal nerves—Diabetes—Neuropathy

E. & Contact Lens 2010;5: 245–248

Corneal confocal microscopy (CCM) is a useful tool for (a) evaluating the integrity of the cornea in conditions such as keratoconus and acanthamoeba and fungal keratitis, (b) assessing the response of the cornea to contact lens wear, and (c) monitoring corneal repair after ophthalmic surgery. The clinical utility of this technique has also been demonstrated for tracking small nerve fiber damage and repair in the cornea after certain forms of corneal refractive surgery when the nerve fiber layer is severed and in a range of peripheral neuropathies including diabetic neuropathy, idiopathic small fiber neuropathy, and Fabry disease.

Two corneal confocal microscopes are currently available: the white light, slit scanning Confoscan 4 (NIDEK, Co., Ltd., Aichi, Japan) and the laser-scanning Heidelberg Retinal Tomograph 3 with Rostock Corneal Module (Heidelberg Engineering GmBH, Dossenheim, Germany). Although these instruments have limited capacity to undertake certain forms of automated or semiautomated assessment of corneal cell structures, neither is equipped to undertake a morphometric analysis of the corneal subbasal nerve plexus. Measurement of this structure is required to facilitate quantitative descriptions of nerve fiber degeneration and repair in conditions such as those outlined above.

Although various authors have reported the application of manual, semiautomated, and fully automated approaches to quantify morphologic aspects of the subbasal nerve plexus, the repeatability of measurement of nerve fibers in the corneal subbasal nerve plexus using CCM has not been reported in the literature. Furthermore, the potential bias resulting from different observers has not been explored. These are important aspects of any diagnostic tool, especially when judgment of an individual observer is involved.

Meijering et al. have noted in some studies that authors have failed to detail their method of nerve fiber analysis, which limits useful comparison between studies. Difficulties in the manual measurement of subbasal nerves have also been discussed in the ophthalmic literature. In view of these concerns and of the imperative to establish the measurement capability of corneal subbasal nerve fiber analysis, we have undertaken a study of the repeatability of a semiautomated technique for quantifying images of corneal nerves captured using the Heidelberg CCM.

MATERIALS AND METHODS

Subjects
Images used in this analysis were obtained from 50 study subjects with type 2 diabetes mellitus with varying severity of neuropathy, who were recruited from the outpatient clinic of the Princess Alexandra Hospital, Brisbane, Australia. Ethical clearance was granted by the Princess Alexandra Hospital and Queens...
land University of Technology Research Ethics Committees, and all subjects participated in the study with informed consent.

**Observers**

The observers were two individual researchers engaged in ongoing studies of corneal markers of diabetic neuropathy at the Queensland University of Technology, Australia. Both were familiar with CCM, although observer 2 was considerably more experienced with the technique.

**Corneal Confocal Microscopy**

All images of the subbasal plexus were captured using the Heidelberg Retinal Tomograph 3 with Rostock Corneal Module. This is a laser-scanning confocal microscope using a 670-nm red wavelength diode laser source. Images produced measure 400 × 400 μm with a digital image size of 384 × 384 pixels. A large drop of a high-viscous eye gel (GenTealEyes; Novartis, North Ryde, NSW, Australia) was placed between the microscope objective and the Perspex "TomoCap" that covered the objective. The eye under examination was anesthetized with a drop of 0.4% benoxinate hydrochloride (Chauvin Pharmaceuticals, Chefaro, United Kingdom), and Viscoatears (Novartis) was applied as a coupling agent between the TomoCap and the cornea. Subjects were asked to fixate on a near target with one eye (generally corresponding to the side of non-dominant hand) while the central cornea of the contralateral eye was examined. At least eight images of the subbasal nerve plexus were captured.

**Corneal Subbasal Nerve Plexus Image Analysis**

Observer 1 chose one CCM image for each subject on the basis of being the first captured, in-focus, high-contrast image of the subbasal nerve plexus. All 50 images were analyzed by this observer using a custom-designed semiautomated nerve analysis software package (CCMetrics13; University of Manchester, Manchester, United Kingdom) that involved tracing the nerves on a Wacom Graphics Tablet (Wacom, Co., Ltd., Saitama, Japan) using a grip pen.

Numerous characteristics of the subbasal nerve plexus can be quantified. For this experiment, we assessed “nerve fiber length” (NFL) because this parameter is considered to be the better test for diagnosing patients with diabetic neuropathy (compared with the parameters “nerve fiber density” and “nerve branch density”) based on received-operator characteristic analysis.14 We classified a nerve fiber in an image as a continuous white line exhibiting a clearly defined path against the grey amorphous background. All nerve fibers and branch offshoots visible in the frame were measured. Any short white line, dot, or mark that was not attached to another nerve fiber (and could, thus, be a keratocyte, Langerhans cell, or image artefact) was not considered as a nerve element and, therefore, was not included in the measurement. Nerve fiber length was expressed as the total length of all nerves and branches in the image (millimeter) per image frame of known dimensions (millimeter per square); thus, units of NFL were millimeter per millimeter square.

After a period of 3 days, all 50 images were analyzed again by the same observer. This process was then repeated by observer 2 using the 50 images originally selected by observer 1. Thus, both observers assessed the same set of 50 images on two occasions separated by at least 3 days.

**Statistical Analysis**

Estimates of NFL were determined for both “individual” subjects (i.e., comparison of the measurements of individual subjects by both observers between successive analysis occasions) and “observer” (comparison of the measurements of images of individual subjects between observers across both analysis occasions), and repeatability was assessed using the intraclass correlation coefficient (ICC). This analysis was undertaken using the “Psychometric,” “ICC.lme,” and “ICC.CI” functions of the “Applied Psychometric Theory” statistical package.13 The correlation between the two “occasions” was also estimated using the statistical package “R.”16 Bland-Altman plots17 were generated to facilitate an appreciation of the extent of between-observer and between-occasion discrepancies and the relation between these discrepancies and the overall magnitude of NFL.

**RESULTS**

The ICC values were 0.95 (95% confidence intervals [CIs]: 0.92–0.97) for individual subjects and 0.95 (95% CIs: 0.74–1.00) for observer. The estimated correlation between the two occasions was 0.96 (95% CI: 0.95–0.98).

Bland-Altman plots of the NFL values obtained on the two measurement occasions by observers 1 and 2 are shown in Figure 1a and b, respectively. For observer 1 (Fig. 1a), a reduced spread of data is associated with lower NFL values. The upward slope of the line in Figure 1a (from left to right) indicates that, for higher mean NFL values, a higher value was assigned to NFL on the first measurement occasion than the second. Similar trends were obtained for observer 2 (Fig. 1b). The overall spread of data was less for observer 2 (standard deviation [SD] = 1.22) than for observer 1 (SD = 2.26).

Bland-Altman plots of the NFL values obtained by the two observers on occasions 1 and 2 are shown in Figure 1c and d, respectively. On occasion 1 (Fig. 1c), a reduced spread of data is associated with lower NFL values. The upward slope of the line in Figure 1c (from left to right) indicates that, for higher mean NFL values, observer 2 assigned a higher NFL value compared with observer 1. Similar trends were observed for occasion 2 (Fig. 1d). The overall spread of data was less on occasion 2 (SD = 1.69) than on occasion 1 (SD = 2.33).

Figure 2A shows an image of the subbasal nerve plexus captured with the CCM that yielded discrepant results between the two observers. The nerve tracing performed by observers 1 and 2 is shown in Figure 2b and c, respectively. (The green dots can be ignored; these indicate designated nerve branches, which did not form part of the analysis reported here.)

**DISCUSSION**

The results of this experiment indicate high between-observer and within-observer repeatability in the measurement of NFL of the human corneal subbasal nerve plexus. The reduced variance in data observed with lower mean NFL values in Figure 1 is to be expected, as a lower number of nerve fibers decreases the possibility of inconsistent scoring and measurement error. Thus, for example, a repeated estimate of NFL in an image containing no nerve fibers or one nerve fiber is likely to be in greater agreement than for a repeated estimate
of NFL in an image containing 20 nerve fibers. The reason why both observers tended to assign a lower NFL value on the second occasion than the first occasion is unclear but may indicate a shift in criteria toward a more conservative approach when identifying nerve fibers on the second occasion.

Between-observer differences in scores assigned on the two separate occasions may be explained by the different skill levels of the two observers. Observer 1 was a visiting researcher to our laboratory who had no previous experience in assessing images of the corneal subbasal layer before this experiment. Observer 2, on the other hand, was an experienced postdoctoral researcher who had been performing confocal microscopy and assessing images of the corneal subbasal layer for 12 months before this experiment. The lower variation of the data points pertaining to the measurements of observer 2 (Fig. 1b) is consistent with the higher level of pertinent experience of this observer compared with that of observer 1 (Fig. 1a). Figure 2 illustrates a large discrepancy in estimating NFL, as evidenced by the greater number of white lines in the confocal image (Fig. 2a) deemed to be true nerves by observer 2 (Fig. 2c) compared with observer 1 (Fig. 2b).

FIG. 1. Relation between differences in NFL vs. NFL mean (a) for observer 1, (b) for observer 2, (c) for occasion 1, and (d) for occasion 2. On each graph, the solid line represents the linear regression and the dotted lines represent the 95% limits of agreement.

FIG. 2. Sample image of the subbasal nerve plexus (a) without any tracings, (b) showing the tracing performed by observer 1, and (c) showing the tracing performed by observer 2. The green dots indicate designated nerve branches, which did not form part of the analysis reported here.
Within-observer differences in the spread of data are possibly explained by a learning effect. By the time the measurements were undertaken on the second occasion, both observers were more experienced, having already undertaken this exercise previously (i.e., the “first” measurement occasion), and undertaking further image analysis tasks during the interim period. It is possible that, on the second measurement occasion, both observers adopted a more purposeful and informed approach to the assessment task, thus narrowing 95% CIs on the second measurement occasion. This effect could be greater in respect of the inexperienced observer (observer 1), who might have become more consistent; thus, the discrepancy in the scores recorded by observer 1 relative to observer 2 would be smaller on the second occasion.

This work documents high repeatability in the semiautomated assessment of images of the human subbasal nerve plexus. The growing interest in assessing this nerve plexus in relation to a variety of ophthalmic applications4–7 and peripheral neuropathic disease6–7 highlights the need for accurate and repeatable quantification of nerve fiber morphology, especially if this measure is to be used in longitudinal studies or indeed to assess the benefits of interventions.

In conclusion, semiautomated measurement of the subbasal nerve fiber layer—in particular measurement of NFL—is shown to be highly repeatable. Repeatability can be enhanced by using more experienced observers. Fully automated systems for analyzing such images, such as those being developed by Scarpa et al.12 and Dabbah et al.,13 hold promise of virtually eliminating remnant errors.

REFERENCES