Automated Software Analysis of Corneal Micrographs for Peripheral Neuropathy

Timothy J. Holmes,[1] Marco Pellegrini,[2] Clayton Miller,[1] Thomas Epplin-Zapf,[1] Sean Larkin,[1] Saverio Luccarelli,[2] and Giovanni Staurenghi[2]

PURPOSE. A relationship has been reported between the presence of peripheral neuropathy and the density and shape of corneal nerve fibers. Peripheral neuropathy is a debilitating condition that arises from many common health problems, and its presence is often confirmed with an invasive clinical test called intramuscular electromyography (EMG). In this study, the possibility of developing an alternative or adjunct test to EMG based on the appearance of nerve fibers in corneal micrographs was explored. Since corneal imaging is virtually noninvasive compared with EMG, such a test may be administered more liberally and frequently, before neuropathy symptoms occur.

METHODS. A software program that automatically traces nerve fibers in corneal micrographs and generates measures based on these traces was implemented. This software was applied to a database of images collected by confocal laser scanning corneal microscopy from diabetic subjects whose levels of neuropathy were measured with EMG and from healthy subjects.

RESULTS. Trends in the nerve fiber density and various measures of shape were calculated and observed, to explore the possibility of using these measures as a clinical tool for corroborating symptoms, confirming an evaluation, or evaluating risk factors for developing neuropathy.

CONCLUSIONS. Preliminary statistical trends show a potential for measuring and observing neuropathy severity or for providing an objective risk measure for a patient's ensuing condition. More work is needed in the development of the measures and in their testing to prove that the measures can be made repeatable in a clinical environment. (*Invest Ophthalmol Vis Sci.* 2010;51:4480–4491) DOI:10.1167/iovs.09-4108

Peripheral neuropathy encompasses disease conditions where the nerve fibers of the peripheral nervous system degenerate. It is a complication brought on by diabetes, aging, alcohol overconsumption, and many other causes.[1] It is among the most common complications of diabetes and is prevalent due to diabetes alone, because diabetes affects ~8% of the North American population,[2] and anyone with diabetes for 25 years has a ~50% chance of having peripheral neuropathy.[3] The most common classification of the condition that occurs in diabetes is distal symmetric polyneuropathy (DSP). The pain and numbness that accompany the condition are often debilitating. It can be dangerous and sometimes life-threatening, because it may affect the autonomous nervous system and thereby may alter physiological functions, most commonly including cardiovascular, gastrointestinal, and genitourinary system functions.

DSP and other forms of polyneuropathy are evaluated mainly from symptoms that include tingling sensations, numbness, and pain, depending on the progression. It may also present as autonomic abnormalities, such as muscle wasting, anhidrosis, difficult food digestion, and other conditions affected by peripheral nerves. The diagnosis is confirmed by using classic neurophysiology techniques, such as electromyography (EMG) and electroneurography; quantitative sensory tests, such as assessment of nerve conduction and study of vibratory thermal and pain sensitivity; and intraepidermal nerve fiber density (INFD) evaluation, which involves a biopsy. These techniques are inadequate for sensitivity in early detection of the disease, for inability to discriminate the type of fiber involved (e.g., A-delta or C fibers), and for their invasiveness. The EMG test cannot discriminate the fiber type and is invasive. A biopsy can discriminate the fiber type but is invasive. Confocal microscopy, which is used in the technique described herein, can discriminate the fiber type and is noninvasive, and so it is an attractive option to develop. The discrimination of type of fiber helps the physician to pinpoint the nature and cause of the neuropathy. Because of the invasiveness of EMG, some forms of electroneurography and INFD tests, there is also the risk of infection, and most important, they are painful and uncomfortable for the patient to undergo, which in turn causes the physician to not prescribe the test liberally. We argue that a test of this sort ought to carry as few complications as possible, so that a physician can routinely prescribe it to rule out or confirm neuropathy as the possible cause of symptoms, and so it can be used to evaluate possible risks for the condition or even possibly as a screening tool.

The EMG test is also commonly used as a screening test. Diabetic children in some countries undergo EMG, typically, once every 1 or 2 years, and older patients (55 and older) undergo EMG yearly. This screening test is critical for detecting early signs of neuropathy. Mainly because of the pain and discomfort, but also because of the invasiveness, it would be good to have an alternative, noninvasive test that accomplishes the same screening objective.

There are three main purposes that motivated this study. One is an early step toward having an alternative or adjunct to EMG, biopsy, and other tests, so that there may be a less invasive screening examination and a less invasive corrobora-

tion of the clinical examination or a complementary further confirmation of the EMG. Another is an early step toward a noninvasive and objective way to evaluate at-risk patients for possible risk indicators or possible early indications that neuropathy may be developing before symptoms occur, so that early preventive treatments can be prescribed. Most commonly, the treatment is of the underlying cause of the condition. A reversible or treatable underlying cause may be identified, such as metabolic, hormonal, or infectious disorders or vitamin deficiencies. With diabetic patients, monitoring and control of insulin levels is tightened and/or patient compliance with insulin therapy is addressed. Third, although there is no widely used treatment today that reverses the underlying neuropathic condition, groups are studying the rudimentary causes[4] that implicate potential treatments, such as vascular endothelial growth factor injection.[5] One proposed nonpharmaceutical treatment that was tested—frequency-modulated electromagnetic neural stimulation[6]—was reported to reverse measurable neuropathic conditions (pain score, tactile perception, and nerve conduction velocity).

The idea of developing software for evaluating the appearance of nerve fibers in the cornea to measure the onset of peripheral neuropathy was implied by others who have shown a correlation between nerve fiber appearance and diabetic neuropathy.[7,8] The tests reported herein are an early step toward automation of this process and establishing a correlation between quantitative measures of nerve fiber shapes and disease progression. Figures 1 and 2 show example images where the nerve fiber densities, shapes, and other appearance measures differ. These densities, shapes, and appearances are being measured by the software.

METHODS

Image Data

Images were collected with a scanning laser ophthalmoscope (HRT II; Heidelberg Engineering, GmbH, Heidelberg, Germany) with a Rostock cornea module, which encompasses a special objective lens that focuses on the cornea.[9]

The study involved 45 patients with type 2 diabetes and, as control subjects, 17 healthy participants. In compliance with the Declaration of Helsinki, written informed consent was obtained from all the patients after an accurate explanation of the purpose of the study and the risks of the procedure. The 45 diabetic patients first underwent a neurologic examination, and EMG was used to rank them as having no, mild, moderate, or severe neuropathy, which corresponds to the rankings denoted as distal symmetric polyneuropathy (DSP) 0, 1, 2, and 3, respectively. Patients who had neuropathy different from DSP, central nervous system diseases, renal failure, autoimmune diseases (e.g., Sjögren's syndrome) and other ocular diseases or previous ocular
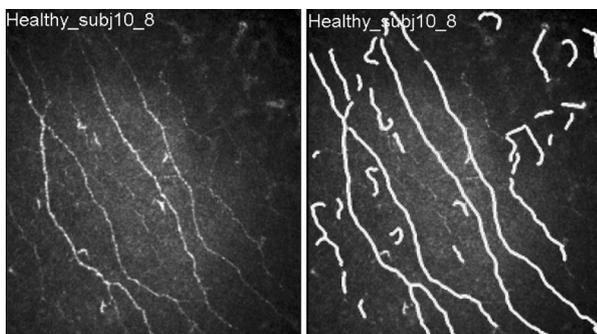
FIGURE 1. *Left*: HRT image of cornea nerve fibers. *Right*: fibers identified by the computer program.
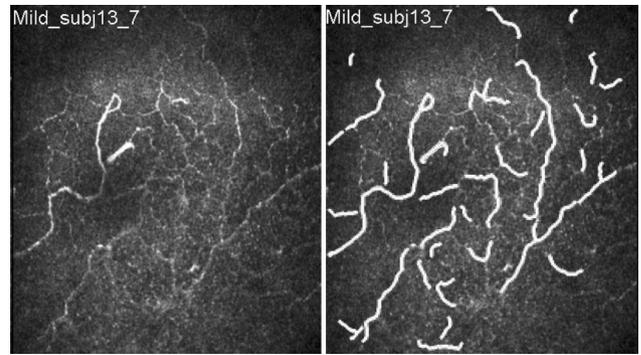
FIGURE 2. Images showing tortuous, lower intensity, and lower contrast fibers.

surgery were excluded from the study. Other information about the subjects included age, sex, years of disease and whether the subjects were symptomatic for neuropathy. Figures 3 and 4 show histograms illustrating subject information that is useful for understanding trends observed in the fiber density and shape measures.

Before the examination began, the cornea was anesthetized with 1 drop of 0.4% oxybuprocaine chlorohydrate (Novesina; Novartis Farma S.p.A, Origgio, Italy). As mentioned earlier, images were then collected with a confocal scanning laser ophthalmoscope (SLO; HRT II; Heidelberg Engineering), on which was mounted the Rostock cornea module, an anterior segment adapter containing a 60× water-immersion objective lens that allows a magnification up to 700× with a transverse resolution of 1 μm. A sterile plastic lens cover (Tomo-cap; Heidelberg Engineering) was mounted over the microscope lens after the application of a drop of contact gel onto the microscope lens.

With the system focusing on the subbasal nerve plexus, many images were collected while the field was moved manually around the center position of the cornea (Fig. 5c). Images were obtained at a large number of positions radial from the center, so that overlapping images were avoided as much as possible. Only one eye for each patient was used, with the selection randomized by coin toss.

A subset of the images for each subject was used for testing trends. This subset was manually selected according to the following three criteria: First, each image in the subset had to have at least one clearly visible nerve fiber. Second, if epithelial cells covered more than one third of the
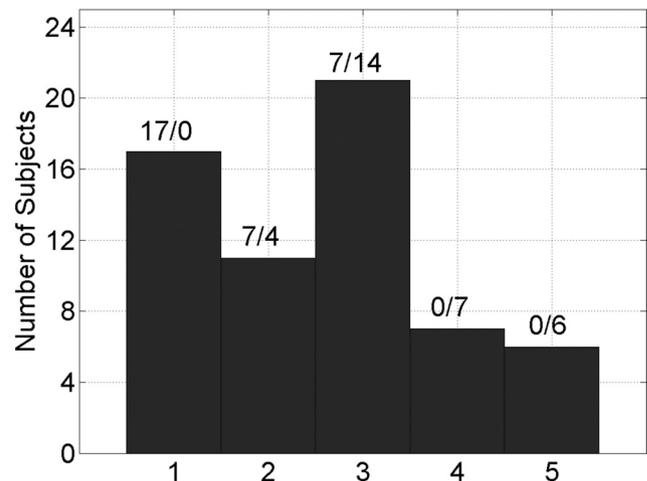
FIGURE 3. Number of subjects versus subject category defined by (1) nondiabetic, (2) diabetic and no neuropathy according to EMG, (3) diabetic and mild according to EMG, (4) diabetic and moderate according to EMG, and (5) diabetic and severe according to EMG. The numbers above each level (e.g., 17/0) indicate the number of asymptomatic and symptomatic subjects, respectively.
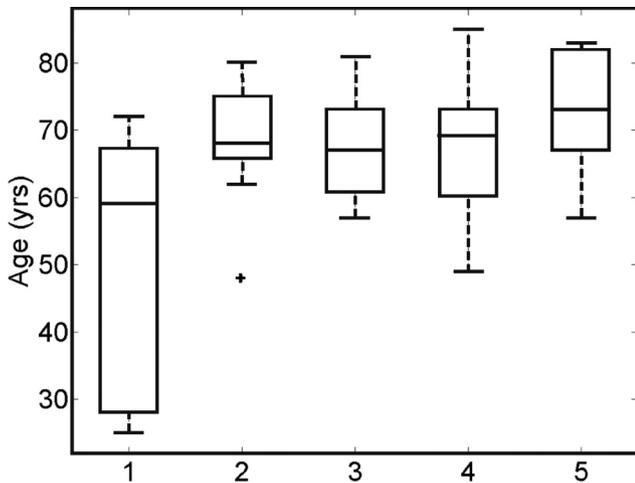
FIGURE 4. Distribution of ages in the subject categories defined in Figure 3.

FIGURE 6. Possible patterns traced by a technician or automated computer program that moves the center of the image field.

image area, that image was discarded. The presence of these cells indicates that the wrong corneal layer is in focus. Third, there could not be too much overlap between images in the subset. If the distance of the shift between a pair of overlapping images was less than half the image width, then the image of lesser quality (evaluated by the viewer based on contrast and clarity of fibers) is discarded. This method creates greater independence between measurements based on different images from the same subject. In the end, while each subject had a different number of images used in the analysis, there was a range of 2 to 24 and an average of ~9 images per subject.

In an effort to eliminate some of the variability in the data due to age, some of the statistical tests were run in subjects who were above specific age thresholds. However, it was not possible to use proper matching according to age or sex because there was an insufficient number of subjects in any age group to provide valid *P*-values in the statistical significance tests.[10,11] The study group contained a large number of volunteers, and for them to be matched according to age and sex, the number of participants would have to increase to an unmanageable size for a first test study. The next step beyond this study will be to perform a rigorously age- and sex-matched study.

One subject was recruited for a study of some of the difficulties with repeatability. One eye was scanned. The acquisition schemes depicted in Figure 5 were performed. Figure 6 shows other schemes that we favored for obtaining good statistics because they would better sample the whole cornea area. We had difficulty in implementing them, however, because of several practical problems that might be overcome with automated collection. The shapes are difficult to realize, partly because it is difficult to have a reference point, partly because the eye moves continually, and partly because it is difficult to see any sort of reference points on the cornea through the camera. A large number (256) of images were collected from the three patterns. Each pattern was repeated once so that two independent collections from each pattern were available for calculating parameters that provide some information on repeatability. Three pairs of such data sets
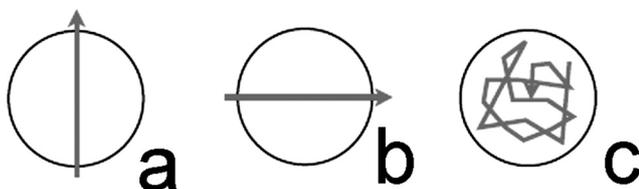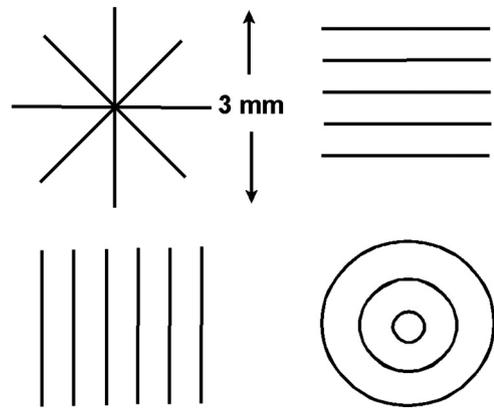
were thereby gathered and given the labels of DU1, DU2, TN1, TN2, R1, and R2, where DU means *down-up*, TN means *temporal-nasal*, and R means *random*. There were 42, 35, 53, 47, 35, and 44 images, respectively, for each of the scans.

In gathering all the data sets, the research adhered to the tenets of the Declaration of Helsinki. Informed consent was obtained from the subjects after explanation of the nature and possible consequences of the study. In addition, approval for the research was approved from the Internal Review Board (IRB) of the Sacco Hospital.

Objectives

The main objectives were (1) to test that the software can automatically reveal trends in the designed nerve fiber measures, presuming that a trend exists. This capability is shown by testing statistical significance against parameters wherein one would expect to see trends or statistical significance, such as age, sex, and visual grading by an ophthalmologist; (2) to see whether the software detects any trend against clinical measures of neuropathy severity, albeit with the limitation that age- and sex-matching precautions could not be taken. Other researchers have already established generally that such trends exist (Pellegrini M, et al. *IOVS* 2008;49:ARVO E-Abstract 2807),[8,12] and the objective in the present study was to see whether the software would show these trends with the automation of nerve fiber detection and the specific measures described later; and (3) to understand the practicalities of usable data acquisition in a clinical environment. To achieve this goal, some parameters reflecting potential repeatability and statistical power were studied.

Automated Measurements

Software was written that automatically identified the fibers as demonstrated in Figure 1. The nerve fibers were detected by using a segmentation and skeletonization algorithm based on ridge map calculation.[13] The program creates an intensity gradient image, which is used to produce a set of labeled pixels in the image called a ridge map. This operation is followed by execution of rules that automatically edit the ridge map to produce a set of labeled fibers. The ridge map is pruned with morphologic operators[14] and further processes follow more rules to find skeletonized branch segments that should be joined together or arranged as having connecting branch points. Counts of branch segments and branch points and measurements of shapes are produced, including those summarized in Table 1 with some of the definitions illustrated in Figures 7 and 8. Later, we refer to these measurements as risk factor measurements.

The definitions of some of these risk factor measurements are according to the equations provided below. A program function returns a list associated with each branch segment of ordered pixel indices (x_n, y_n). This ordered list traces the branch segment's skeleton. A spatial low-pass filter was applied to the function that this list

FIGURE 5. Simple patterns that could be traced by a technician or automated computer program. In this study, the traces were manual. The drawing in (**c**) represents randomly chosen locations.

TABLE 1. Candidate Risk Factor Measurements

Measurement	Description
Number of branch points, N_b	Number of branch points in the image.
Number of branch segments, N_S	Number of branch segments in the image.
Branch segments per branch point N_S/N_b	Number of branch segments divided by number of branch points.
Curvature tortuosity 1, T_{C1}	The tortuosity defined in equation 5 and illustrated in Figure 8, left. This parameter is based on the mathematical curvature of the skeleton of the branch segment.
Curvature tortuosity 2, T_{C2}	The tortuosity defined in equation 8 and derived from T_{C1}.
Length, L	Path length of the branch segment.
Length density, D_L	Sum of lengths of all branches divided by the area as defined in equation 7. This is, de facto, a measure of total length over the image field. The study could have equivalently recorded the lengths only, because the area divisor is the same for all images, and so dividing by the area has no consequence on the trends.
Length ratio tortuosity, T_L	The tortuosity measurement illustrated on the right of Figure 8. It is the path length of a fiber divided by the distance between its end points. This measure has been called the arch length over length ratio.[15]
Signal to background, S/B	Ratio of the mean intensity value within pixels on the fiber to the mean intensity value within surrounding pixels.
Intensity variance, σ^2	Variance of intensity values within pixels on the fiber.
Width, W	Total fiber area divided by length.

represents to remove rapid shifts from the pixel quantization of the (x_n, y_n) positions. The resulting filtered sequence is then processed according to the following formulas to form several measurements:

$$\Delta x_n = x_n - x_{n-1} \tag{1}$$

$$\Delta y_n = y_n - y_{n-1} \tag{2}$$

$$\Delta^2 x_n = \Delta x_n - \Delta x_{n-1} \tag{3}$$

$$\Delta^2 y_n = \Delta y_n - \Delta y_{n-1} \tag{4}$$

and

$$T_{C1} = \frac{1}{L} \sum_n \left| \frac{\Delta x_n \Delta^2 y_n - \Delta^2 x_n \Delta y_n}{(\Delta x_n^2 + \Delta y_n^2)^{3/2}} \right| \tag{5}$$

where L is the length of the segment defined by

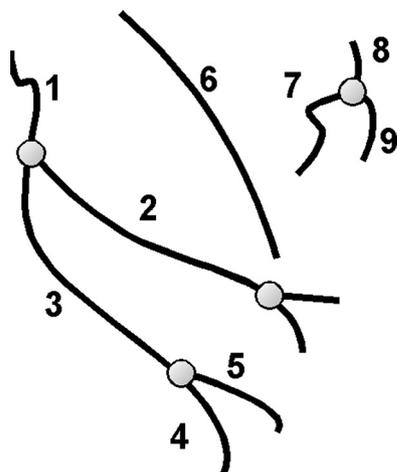$$L = \sum_n \sqrt{\Delta x_n^2 - \Delta y_n^2}. \tag{6}$$

The length density, D_L, is defined according to

$$D_L = \frac{\sum_k L_k}{\sum_i A_i} \tag{7}$$

where the index k refers to the branch segment being considered and the subscript i refers to the ith image used for the subject being measured.

The T_{C1} parameter is called *curvature tortuosity 1* because it is based on the mathematical curvature of the skeleton of the branch segment.[15] This tortuosity is identical with the $Tc(C)$ measure that is defined in Hart et al.,[16] except that here the term is divided by L which weights it inversely to the length of the branch segment. We also have a *curvature tortuosity 2* (T_{C2}) defined according to

$$T_{C2} = \frac{\sum_k T_{C1k} L_k}{\sum_k L_k} \tag{8}$$

where T_{C1k} is the T_{C1} measured for branch segment k. A third measure, called the length-ratio tortuosity,[15] is defined according to

$$T_L = L/L_e \tag{9}$$

FIGURE 7. Definitions of terms for specifying measurements. The objects enumerated from 1 to 9 are branch segments. *Circles*: branch points.

FIGURE 8. The difference between curvature tortuosity and length-ratio tortuosity. The branch segment on the *left* has higher curvature tortuosity and the branch segment on the *right* has a higher length-ratio tortuosity.

where L_e is the distance between the two endpoints of the branch segment.

Ophthalmologists in our group described to us the characteristics that they expected to trend with neuropathy severity, based on their experience. Most of the risk factor measures were designed to reflect these characteristics. Fiber length, count, and branch points, they said, all are expected to lessen with severity. Tortuosity should increase, according to their expectation, whereas beading and contrast should worsen. The variance calculation σ^2 was designed to reflect the beading, since as beading increases, the variation of intensities in the fiber appear to increase. The signal-to-background (S/B) was designed to capture the changes in contrast.

The software performs all the fiber identification and candidate risk factor measurements automatically. There is an operator intervention where thresholds of fiber lengths and other measurements can be set to control rejection. For example, in Figure 9 it is explained that all fibers with a length <25 μm and having an S/B measurement of <1.25 were rejected, but other than the selection of these thresholds, this rejection process is automated. An obvious idea is to edit the traces by manually drawing missing branches, erasing spurious branches or branch segments, or joining fragmented sections with a point-and-click drawing interface. With the data presented, there was no such operator editing. The software makes errors, as is evident by the missed and spurious fiber identifications in Figures 1 and 2. These errors ought not to be overlooked, and future improved versions of the software algorithms ought to provide more accurate traces. Ways of doing so are suggested later.

It is expected that researchers and certain clinics will want to edit the fiber traces. Clinically, such an editing feature may be desired for visual inspection and evaluation of the images, aside from statistics and quantitative measurements and for communicating the condition to the patient. Even so, it may be better overall for the statistics and the derivation of the quantitative risk factor measures to accept the automation mistakes and in the meanwhile to work on improved automated detection of the fibers, than to edit them manually. Editing introduces nonrepeatable human error that adds to statistical variability. Our experience with other projects[17,18] has shown that perfect nerve fiber tracing is not possible, either manually or automatically and that attempting to edit the fibers is laborious. Even manual traces, performed by an ophthalmologist and shown in Figure 10, will contain mistakes, equally severe, and will not be perfect, but will have the added variability between ophthalmologists and within the same ophthalmologist at different times. Future improved versions of algorithms will strive to eliminate such automated errors and will have the editing feature as well.

One of the problems that prevents perfect decisions about what is a branch, whether the computer or the person decides, is that every image set contains cases with reflectance, contrast, and branch continuity that are borderline for being counted. The arrows on the top left

of Figure 10 highlight one such example. There are other examples apparent in this image. Those that are highlighted could be considered mistakes, although it is not clear that they are. Both the computer and the person make similar types of mistakes on such decisions, and in such cases the correct decision is not clear, which adds to the variability. However, the person (or persons) will not make the same decision every time she is presented with the same situation. This added change-of-mind variability does not happen with the computer. Mistakes by the computer are a source of variability, but the nonrepeatability of the decisions is not.

Another argument against editing fragmented branches is that they reflect the degradation of the fibers caused by neuropathy (of course, some fragmentations could also be due to the branch's pathway in and out of the confocal plane). Perhaps, it is important to capture the quantitative measure of these fragmentations and correlate them with the neuropathic condition. If so, it may be better to leave them fragmented, counted, and measured as smaller branch segments. On the other hand, an even more elegant measure would be a number that reflects the nerve fiber breaks per nerve fiber, and this number would require both the fragmented branches and their correct association with a corrected, unfragmented branch. Considering these and similar arguments, research end-users of this software may want the capability of editing the decisions as to what are branches and of editing the branch tracings that affect the end measurements. Of particular interest, and the easiest and potentially most relevant feature needed, is the capability of joining fragmented branches. This type of edit is conducive to automation, although the current software version does not do so yet. A simple automated postediting scheme could detect that adjacent branch segments are close to one another and point in generally the same direction and could make the automated decision to join them. Of course, such a feature must be investigated, because one can envision many types of additional mistakes that such a scheme could introduce. The eventual commercial software product ought to provide this capability, whether it be an automatic postprocess, a manual editing feature, or both. Manual edits that rejoin fragmented branches would be the most benign of edits because the computer has already decided on where the branches are, and the user can simply join them. This decision will vary less than decisions about what constitutes a branch.
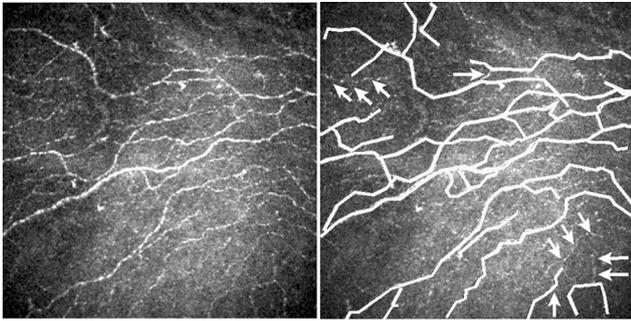
RESULTS

Trends

With each subject, the software processed every nonoverlapping image that was selected, as described.

Figure 9 demonstrates trends with age. Similar trends were seen when these plots were repeated for diabetics only and for
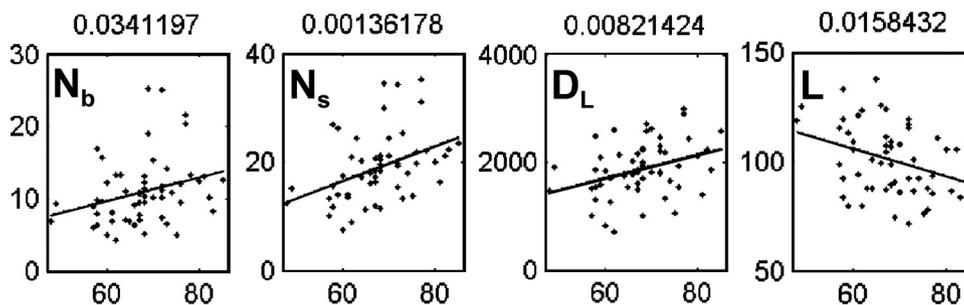
FIGURE 9. Scatterplots and linear regression of measures versus age in all diabetic and nondiabetic subjects age 48 and above. Each point represents a volunteer subject. The *P*-value is shown above each plot and is based on the significance test used for regression analysis and outlined in Sec. 10.3 of Reference 10. Branch segments with $L < 25$ μm were excluded. Branch segments with $S/B < 1.25$ were also excluded. For N_b, N_s, and D_L, each scatter-plot point was calculated by first calculating the measurement in each image and then taking the average of that measurement over the included images of that subject. For L, each scatter-plot point was calculated by summing the lengths of every included branch segment throughout all included images and then dividing by the number of included branch segments throughout all images.

FIGURE 10. An example of a manual trace. Note the wide range of contrast in the fibers, which makes it difficult to decide what makes up a branch. *Arrows*: examples of so called branches that result in different decisions, depending on the operator and depending on when the same operator made the decision.

nondiabetics only, but the resulting *P*-values were not as strong. Figure 11 demonstrates differences between the sexes.

Figure 12 demonstrates trends against graded images. An ophthalmologist (MP) viewed all the images in the database for chosen subjects and graded them as one of three levels: mild, moderate, or severe. The 27 subjects for this test were chosen randomly from all groups in the database. From each subject, the image to grade was chosen randomly as well, except that poor-quality images (those with poor contrast, excessive noise, or no fibers) were avoided. The ophthalmologist did not follow any specific instructions except to rank them as mild, moderate, or severe. Afterward, he explained that he considered the appearance of number of fibers (without actually counting), the appearance of their collective lengths (without actually measuring), and their tortuous appearance.

To test trends with neuropathy progression, there are three tags that were candidates to correlate or trend with the candidate risk factor measures: the diabetic/nondiabetic tag, the asymptomatic/symptomatic tag, and the no-neuropathy/mild/moderate/severe EMG test output that was performed in all diabetic patients. Figure 13 shows the results of a test in which the subjects were grouped according to the following criteria:

Group A includes subjects who met either of the following conditions: all nondiabetics older than 55 years or diabetics with no neuropathy according to EMG and no clinical symptoms, and older than 55.

Group B includes subjects who met the following conditions; diabetics who had an EMG evaluation of mild neuropathy, were clinically symptomatic, and were older than 55 years.

The measures were averaged by including only the longest three branch segments in each image. The *P*-value shown above each plot was based on a paired *t*-test. Averaging for each subject was performed by measuring every included branch segment among all included images and then taking the average over that whole population of branch segments.

Figure 14 shows the results of a test in which the subjects were grouped according to the following criteria:

Group C was all nondiabetics of all ages.

Group D was all diabetics with an EMG evaluation of no neuropathy (included both sexes and both asymptomatic and symptomatic disease).

Group E was all diabetics with EMG evaluation of mild (included both sexes and both asymptomatic and symptomatic disease).

The measures were accomplished by including only the longest three fibers in each image. All other fibers were not included in the calculation. The *P*-value shown above each plot is based on ANOVA.[10,11]

The measures L, T_{C1}, T_L, S/B, σ^2, and W were averaged for the box plots and scatterplots of Figures 9 and 11 to 14 by averaging them over every branch segment within all included images of the subject. The measures N_b and N_s were averaged for these figures by calculating them for every image of the subject and then averaging the value over these images. The length density D_L, as shown by equation 7, was calculated by adding all the lengths of included branch segments of a subject and dividing by the sum of the areas (micrometers squared) of all the images of that subject. This measure is equivalent to calculating the aggregate length of the fibers in each image, and then averaging that aggregate length over the images, since
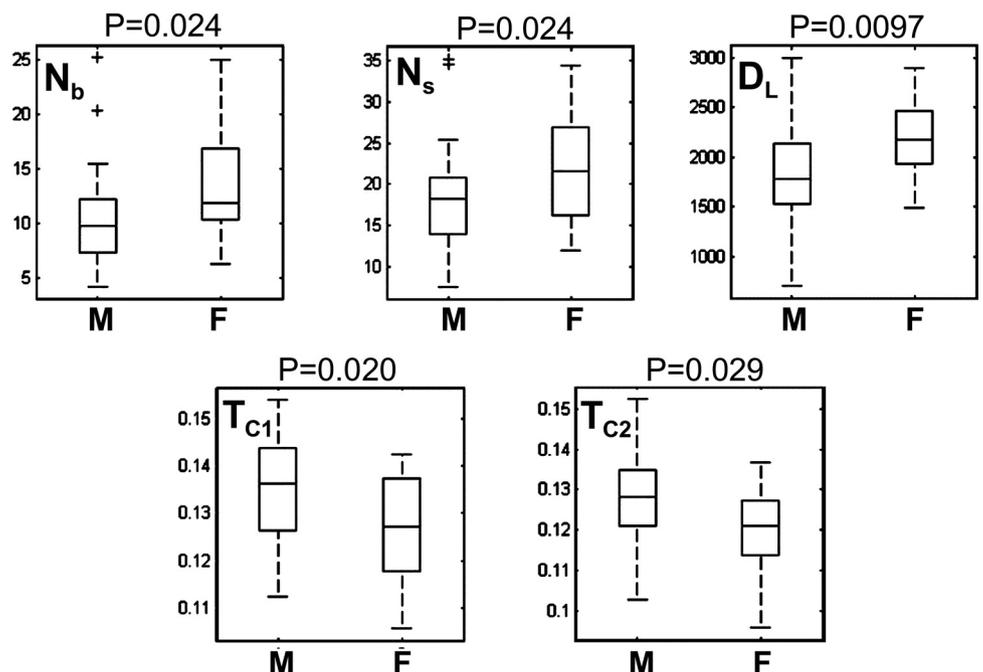
FIGURE 11. Paired *t*-test of measures separated by sex. Ages 55 and under were excluded. Branch segments with $L < 25$ μm were excluded. Branch segments with $S/B < 1.25$ were also excluded. The sample point representing each subject, for N_b, N_s, and D_L, was calculated by calculating the measurement in each included image of that subject and then taking the average of that measurement over all the included images of the subject. The sample point representing each subject, for T_{C1}, was calculated by measuring every included branch segment among all included images and then taking the average over the whole population of included branch segments. The sample point representing each subject, for T_{C2}, was calculated by having the summations in both the numerator and denominator of equation 8 being taken over all branch segments throughout all included images of the subject.
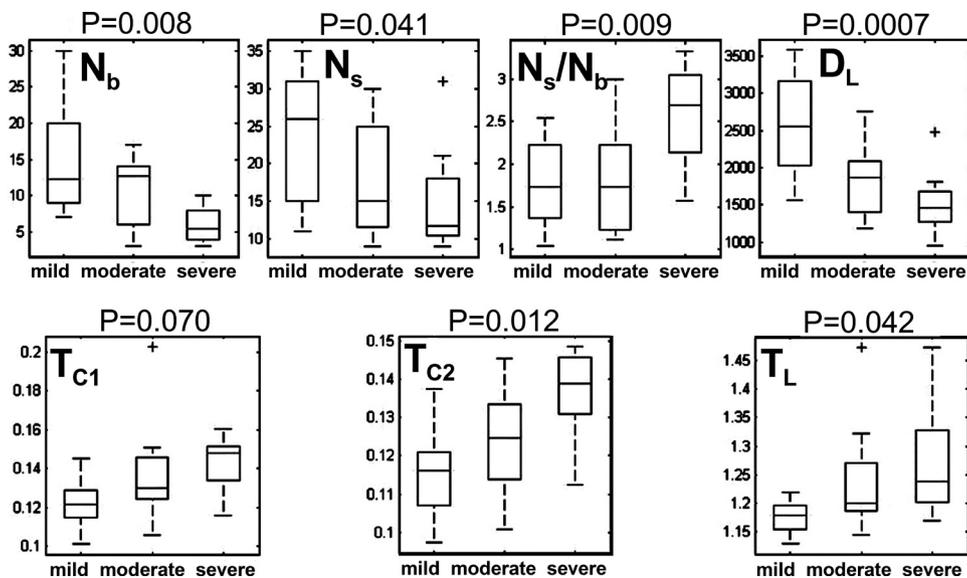
FIGURE 12. Measure trends with grading by an ophthalmologist. The *P*-value was based on the ANOVA statistical significance test.[10,11] Branch segments with $L < 25$ μm were excluded. Branch segments with $S/B < 1.25$ were also excluded. The sample point representing each subject, for N_b, N_s, D_L, T_{C1}, and T_{C2}, was calculated according to the description in the Figure 11 caption. The sample representing each subject, for N_s/N_b, was calculated by first calculating N_s/N_b for every included image of the subject and then averaging this ratio over these images. The sample point representing each subject, for T_L, was calculated by measuring T_L for each branch segment throughout all included branch segments of the subject and then taking the average over the whole population of branch segments.

division by the area simply scales the number. The *k* index of T_{C2} in equation 8 implies that the calculation is performed over all included branches and all included images. The branch-segment to branch-point ratio was calculated by finding the averages of both N_s and N_b and then taking the ratio.

Repeatability Considerations

There are two immediate practical concerns that drive the need for repeatability. First, it is important that the same evaluation be arrived at with a patient if he is examined twice
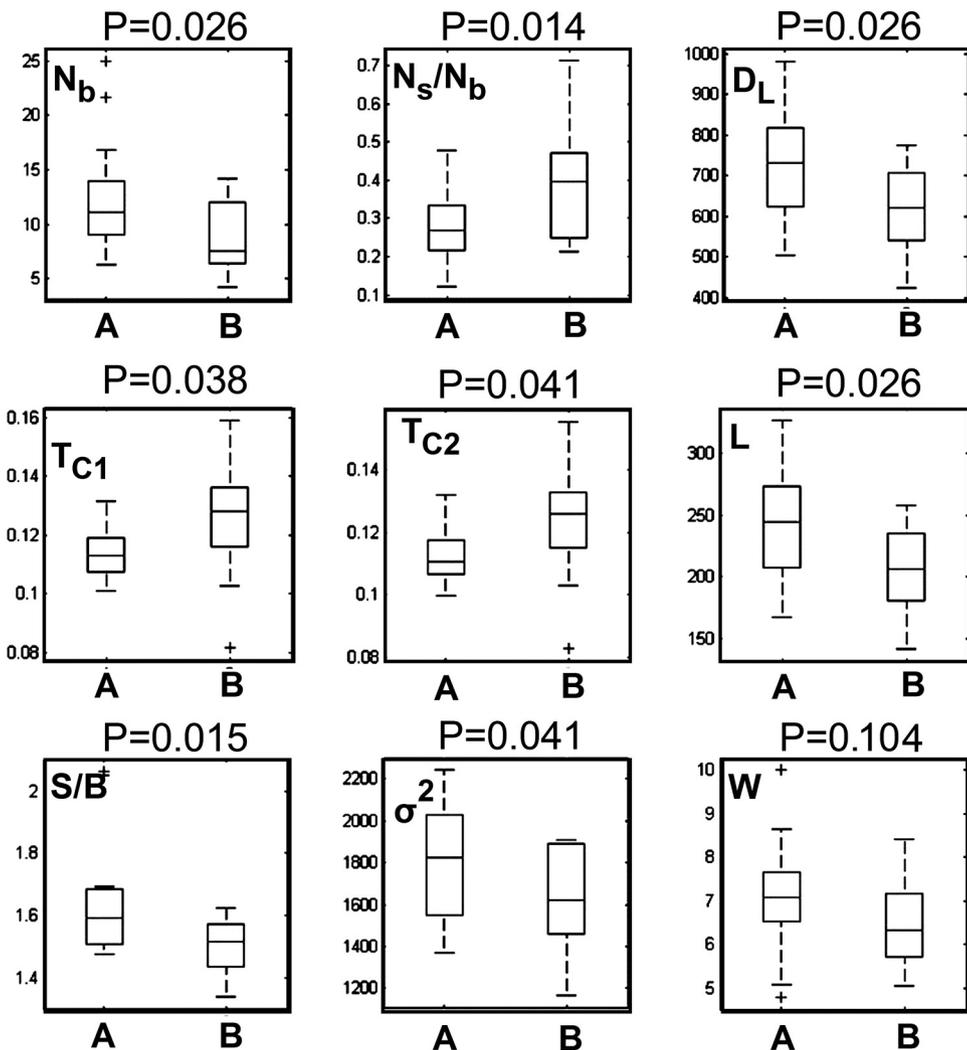
FIGURE 13. Measures versus clinical evaluation groups according to the following criteria. Group A includes subjects who met either of the following conditions: (1) all nondiabetics older than 55 years, (2) diabetics with no neuropathy according to EMG and no clinical symptoms, and older than 55. Group B includes subjects who met the following conditions: diabetics who had an EMG evaluation of mild neuropathy, were clinically symptomatic, and were older than 55 years. The criterion for including fibers in the measurements was that only the longest three fibers of each image were included. No *S/B* criterion was used as with the prior figures. The *P*-value shown above each plot was based on a paired *t*-test. The sample point representing each subject, for N_b, N_s, D_L, N_s/N_b, T_{C1}, and T_{C2}, was calculated according to the descriptions in the Figures 11 and 12 captions. The sample point representing each subject, for *L*, *S/B*, σ^2, and *W*, was calculated by measuring these parameters for each branch segment throughout all included branch segments of the subject and then taking the average over the whole population of branch segments.

FIGURE 14. Measures versus clinical evaluation groups according to the following criteria: Group C was all nondiabetics of all ages. Group D was all diabetics with an EMG evaluation of no neuropathy (included both sexes and both asymptomatic and symptomatic disease). Group E was all diabetics with EMG evaluation of mild (included both sexes and both asymptomatic and symptomatic disease). The measures were accomplished by including only the longest three fibers in each image. All other fibers were not included in the calculation. No *S/B* criterion was used to filter fibers as was done for previous figures. The *P*-value shown above each plot is based on ANOVA. Refer to earlier figure captions (Figures 9–13) for an explanation of the calculation of each sample point that represents each subject in these plots.
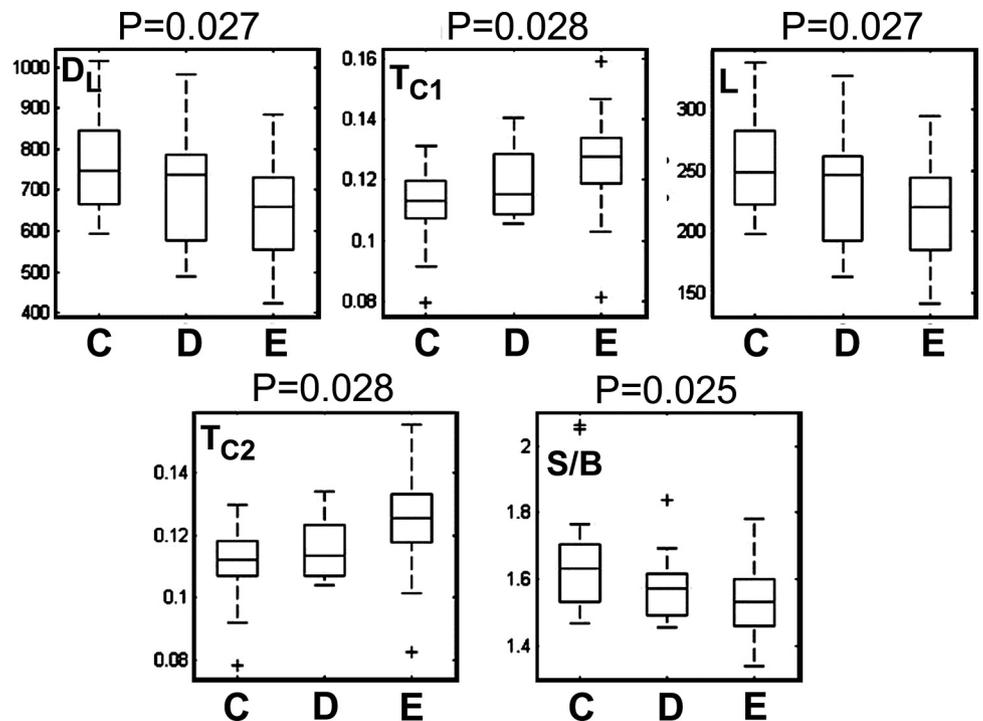
under the same conditions. Second, it is important to reliably detect a progression of any of these potential measures, and the more sensitivity in doing so the better. One way of characterizing this sensitivity is by the effect size under realistic conditions.

An early test of repeatability is summarized in Table 2. The risk factor parameter values were recalculated for the image sets taken under identical conditions (i.e., DU1 versus DU2, TN1 versus TN2 and R1 versus R2). The null hypothesis was tested for each of these pairs. For the paired *t*-test, the measurement calculated for an individual fiber made up the sample point L, T_{C1}, S/B, σ^2, T_L, and W. The measurement averaged over an image made up the sample point for N_b, N_s, D_L, and T_{C2}. For T_{C2}, this average followed equation 8. A *P*-value greater than 0.05 implies that the difference in means is statistically insignificant. Ideally, we would like every entry to have high *P*-values.

Most of the tests showed the desired statistical insignificance, but 6 of the 30 tests did not. Four of the measures (N_b, T_{C2}, L, and T_L) showed statistical insignificance in all three tests. For future clinical usage, we think it is important to collect the baseline and follow-up data in the same way. For

example, the worst repeatability occurred when comparing the means between DU1 and R1 (not shown in Table 2). In this case, all 10 of the measures had a statistically significant difference. A definitive conclusion about repeatability cannot be arrived at from these tests, but they at least provide some insight, and they imply that, in the clinic, it will be important to define and adhere to collection patterns. We think that other collection patterns and measures could improve repeatability. A thorough repeatability test necessitates compiling a well-defined database that represents a longitudinal study including healthy control subjects, to ensure no change, and subjects who are expected to show a change.

To roughly estimate a possible realistic sensitivity, we used the TN1, TN2, DU1, DU2, R1, and R2 data sets as examples for estimating a possible effect size and concomitant percentage change in parameters that could be detected. The standard methods of estimating effect size from desired statistical power were used.[11] Tables 3 and 4 summarize these calculations, which shows the estimated percentage change needed for a statistical power of 0.8, estimated from the DU1 data set. The DU1 set was chosen arbitrarily. Table 3 summarizes the param-

TABLE 2. Calculation of Repeatability

Measure	TN1 vs. TN2		DU1 vs. DU2		R1 vs. R2	
	% Difference	*P*	% Difference	*P*	% Difference	*P*
N_b	13.2	0.348	−10.3	0.516	−21.3	0.058
N_s	24.6	0.016*	1.8	0.879	−3.3	0.695
D_L	25.1	0.012*	5.9	0.546	−5.5	0.341
T_{C2}	0.72	0.844	−3.4	0.25	18.0	0.054
L	2.6	0.649	13.4	0.062	−6.0	0.375
T_{C1}	2.4	0.602	−3.3	0.263	10.9	0.002*
S/B	4.1	<0.001*	−0.83	0.568	−2.8	0.077
σ^2	3.3	0.469	20.0	0.001*	−9.7	0.072
T_L	−16.2	0.285	−1.89	0.543	3.5	0.081
W	10.7	0.021*	13.1	0.027*	2.2	0.568

* Statistically significant.

TABLE 3. Percentage Change Needed for a Statistical Power of 0.8, Estimated from the DU1 Data Set. These Are Measures Whose Means Are Calculated over All Included Fibers in the Image Set.

Measure	Estimated Mean	Estimated SD	Required Change in the Mean	Required Percentage Change
L	93.17	70.76	11.86	12.7
T_{C1}	0.139	0.048	0.008	5.8
S/B	1.48	0.226	0.038	2.6
σ^2	1198.8	770.14	129.06	10.8
W	5.84	3.67	0.616	10.5

eters that are averaged over all included fibers. Parameters needed for the Table 3 estimates included the number of fibers (441, taken from the number of fibers automatically identified in DU1), the desired null-hypothesis P-value (0.05), and the desired statistical power (0.8). The required effect size (percentage change over the SD) for this P-value, statistical power, and number of fibers is 0.168. Table 4 summarizes the parameters that are averaged for each included image and then averaged over the images. Parameters needed for the Table 4 estimates include the number of images (42, which is the number of images used for DU1), desired null-hypothesis P-value (0.05), and desired statistical power (0.8). The estimates shown for the mean, standard deviation, and required change in the mean were also calculated from the DU1 data set. The required change in the mean is calculated by multiplying the estimated standard deviation by the required effect size. The required percentage change in the mean is calculated by dividing the required change in the mean by the estimated mean.

Table 5 provides a summary of how the numbers calculated in Tables 3 and 4 vary over all six of the data sets (DU1, DU2, TN1, TN2, R1, and R2) by showing their ranges over these data sets. Also shown in the second and third columns of Table 5 are estimates of a possible typical percentage change that can be expect from a patient when the severity progresses. These expected percentage changes cannot be known without a complete study, but at least they may serve as first estimates based on the representations of ranges provided in Figures 13 and 14. The percentages in these two columns were estimated by calculating the percentage differences between the median values shown in the box-and-whisker plots of Figures 13 and 14, respectively. For the Discussion these calculated percentages serve as a surrogate for estimating what might be possible in a real situation. They do not report what percentages will actually be reached, because more data would be needed to determine those statistics.

DISCUSSION

Figures 9 and 11 to 14 show that the software is capable of showing trends in measures against parameters where a trend would be expected. Although there is no prior determination that corneal nerve fiber shapes are, on average, different between males and females, it is not surprising to find. Other researchers have documented differences in nerve fiber size, shape, number, and function between males and females in animals and humans, in a variety of tissues.[19–24] This includes the documentation of increased fiber counts (N_b, Fig. 11) in cadaveric skin specimens of human females compared with those in males.[23]

The linear regression plots versus age shown in the first three plots of Figure 9 trend in the opposite direction from what was expected, but are consistent with those in a published study in which nerve fiber length densities were measured in skin biopsies of diabetic patients.[25] We expected the

number of branch segments, branch points, and length density to decrease with age, since we thought that they were indicators of fiber health and further assumed that general health declines with age. On the other hand, we confirmed these trends with the visual appearances of the fibers in the images, and these trends appear to be real. A possible explanation is that the nerve fibers are in continual flux,[26–28] and so it is believable that the fibers could generate continually, and the number could, on average, increase over time and thereby over the age of the subject. Before drawing conclusions based on this finding, a regression analysis on a wider range of population has to be performed. Data sets of subjects of age <55 years were not included in Figure 9, because there were no diabetic subjects in this lower age category, and there was a good mix of both diabetics and nondiabetics in the range of 55 years and above. A further corroboration of this trend is provided in Figure 1 in Reinisch et al.,[25] where a regression analysis shows an increase in fiber length density (length per unit area) of the subepidermal nerve plexus in diabetic subjects on the order of 50% between the ages of 50 and 90, similar to that shown in Figure 9. The main difference in our data, other than being extracted from different tissue, is that our pool of subjects was a mixture of both diabetics and nondiabetics (although most were diabetic), whereas the data in Reinisch et al.[25] are from diabetic persons only.

Figures 13 and 14 offer hope that it may become feasible to use risk factor measurements of this sort. A determination of the usefulness of these candidate risk factor measures, beyond just showing feasibility, requires several more studies. A longitudinal study is needed, in a group of control subjects and a group of diabetic subjects, to verify that the software will accurately follow progressing conditions. Future studies should include age- and sex-matching as well as other precautions, to remove variability in the data. One ultimate clinical application is to confirm the likelihood of neuropathy after the information has been gathered from clinical neurologic or other examinations, such as EMG. Another is to periodically perform preventive examinations, or risk screening, of diabetics and other at-risk patients, as is done now with EMG and electroneurography.

EMG is used for this preventive purpose, but it is relatively invasive and causes anxiety and discomfort to the patient. The confocal microscopy that collected the measures in this study would not have this problem, because it is essentially noninvasive. The most invasive factors are that an optical-coupling gel is applied to the cornea, which carries a small risk of infection. Light is injected into the eye, but the amount of light is well below safety limits.

In any of the expected applications, a baseline examination is needed. The alternative of establishing a normative database of risk factor measurements, to know progression states from an absolute measure, does not seem likely, considering the variance indicated by the box-and-whisker plots in Figures 13 and 14. There will be too much variability in the measurements

TABLE 4. Percent Change Needed for Statistical Power of 0.8, Estimated from the DU1 Data Set. These Are Measures Whose Means Are First Calculated per Image and Then Averaged over All Included Images in the Set.

Measure	Estimated Mean	Estimated SD	Required Change in the Mean	Required Percentage Change
N_b	5.10	3.40	1.86	36.5
N_s/N_b	2.85	2.14	1.17	41.2
D_L	978.2	443.9	242.9	24.8
T_{C2}	0.13	0.017	0.009	7.0

TABLE 5. Ranges of Required Percentage Changes

Measure	Range of Required Percentage Change	Percentage Change Estimated from the Difference between Categories A and B in Figure 13	Percent change Estimated from the Difference between Categories C and E in Figure 14
L	11.8–14.0	15.3	$\approx 15\%$
T_{C1}	5.1–24.8	13.1	$\approx 12\%$
T_{C2}	6.5–24.2	13.7	11.9
S/B	2.3–3.0	4.9	6.24
σ^2	9.3–10.8	11.2	NA
W	8.9–11.2	10.5	NA
N_b	28.5–47.7	32.2	NA
N_s/N_b	24.0–41.2	47.7	NA
D_L	14.1–24.8	15.3	11.8

Calculated by repeating the calculations in Tables 3 and 4 for all the DU1, DU2, TN1, TN2, R1, and R2 data sets.

from subject to subject to establish nonoverlapping ranges of the indices that classify disease state. With a baseline examination, it is hopeful that changes in the measures compared with a patient's own baseline values will be detectable. We envision establishing ranges of these changes from baseline for indicating mild to severe classifications.

Although the topic of this article is primarily focused on diabetic neuropathy, there are other applications to the automated measurement of nerve fibers in the cornea. Such applications include peripheral neuropathy as a complication due to other conditions, including genetic diseases[29] and many inflammatory diseases including lupus erythematosus and Sjögren's syndrome, vitamin deficiencies, HIV (Sabato L, et al. *IOVS* 2008;49:ARVO E-Abstract 2804), chemotherapy, and others. As explained in Scarpa et al.,[12] corneal surgical interventions, including LASIK, photorefractive keratectomy, and transplantation disrupt the integrity of the nerve fibers, and regeneration of these fibers occurs after surgery (Midena E, et al. *IOVS* 2008;49:ARVO E-Abstract 2261). The capability of automatically identifying nerve fibers and measuring them may provide an objective means of evaluating reinnervation after these surgeries.

There are several sources of potential variability that may have degraded the trends in Figures 13 and 14. The image fields were located randomly, as described earlier, and it is known that the nerve fiber appearances vary widely across the cornea.[26–28] It is possible that, in some subjects, good sampling over all regions of the cornea was not obtained. Considering that the nerve fiber numbers and shapes change across the cornea positions, especially radially, a better collection scheme would use a pattern like one of those illustrated in Figures 5a, 5b, and 6, and would collect a whole stack (i.e., axially) at each location. The ideal would be to image exactly the same fields between baseline and follow-up visits and thereby ensure that the same fibers are measured. One may think of ways of doing so, but there is one problem that is impossible to circumvent. The nerve fibers are not stationary. They are in continual flux, moving and changing shape from one visit to the other.[26–28] The best hope is to strive for stationary statistics by using schemes like those shown in Figures 5 and 6. The schemes in Figure 6 are better for more thorough sampling of the cornea, but work is needed on the software that automates the collection of these patterns. The schemes are difficult to execute manually because it is hard to determine a reference point on the cornea, and it requires a relatively long time to collect. With a contact microscope, the schemes in Figure 5 are more practical because they require shorter contact time with the cornea. The scheme of Figure 5c was used for collecting the

data pools of Figures 13 and 14. All three of the schemes in Figure 5 were used to examine repeatability.

Criteria for selecting which image in the stack to measure must be designed, such as finding the one with the highest contrast, longest fibers, or largest number of fibers. Criteria for selecting which fibers to include in the statistics of a risk factor measurement must be determined. For example, we learned that including just the longest three fibers from each image produced the most significant trends shown in Figures 13 and 14. One can think of other filtering schemes that may be even more effective at eliminating variability.

Age and sex matching were generally not used, although some of the tests (Figs. 9, 11, 13) excluded participants under a specific age. Figures 9 and 11 indicate that the nerve fiber appearances change with age and sex, so matching of these two parameters would be important and helpful in designing a useful clinical tool.

Most of the subjects who were known to have neuropathy were being treated for the condition and therefore the mild, moderate, and severe rankings may not have represented a wide enough range of severity. Using treated subjects probably contributed to the difficulty in trending the measurements against these rankings by themselves. Second, some subjects who may have a relatively advanced progression may appear asymptomatic because they are receiving medication to suppress the symptoms, making it difficult to trend against the asymptomatic/symptomatic tag by itself. There were few moderate or severe disease subjects (7 and 6, respectively), which further made it difficult to trend against these categories.

As mentioned, the software makes some mistakes in identifying the fibers. Spurious and missing fibers can be seen in Figures 1 and 2. Realistically, the philosophy behind improving the software's capability is to continually strive for improvement in eliminating mistakes and to understand that mistakes will never be completely eliminated. Perfection in tracing the fibers without error can be approached and should be a goal, but such perfection will never be achieved. Such has been our experience in other projects.[17,18] These mistakes probably do not contribute significantly to difficulties in establishing trends, because such mistakes are repeatable and systematic and because the other sources of variability are obvious and significant. It is an obvious suggestion to place a human operator in the process by editing the nerve tracings as discussed earlier in the Automated Measurements section, but doing so may reduce repeatability and, as a result, the capability of detecting trends and thereby recognizing progression of a condition.

As shown in Figures 3 and 4, within the relatively wide range of subjects who were examined, there was little uniformity in sampling various age groups, and this factor must have adversely affected the ability to accurately determine trends. Therefore uniform and fair sampling of diabetics, nondiabetics, sex, age, image fields, and many other factors were not represented in these data sets.

Conclusions

This article demonstrates a step toward a potential novel technology for measuring and observing potentially important risk factors of peripheral neuropathy. Probably the most compelling aspect of the work is that the fiber identifications and their measures are automated, which eliminates the manual labor, tedium, and human operator variability. If proven to be valuable, it is envisioned to be used as an adjunct to other clinical examinations and measurements, such as EMG and electroneurography tests. It is less invasive than other tests that are currently used for similar purposes, such as the EMG and skin biopsy. On the other hand, because it requires images of cornea tissue from an SLO, it will have any risks associated with capturing cornea images with an SLO that has a lens in contact with the cornea. Risks are infection or injury to cornea tissue due to the contact. The Confoscan 4 model (Nidek Co., Ltd., Gamagori, Japan) uses a noncontact objective lens, and so it does not have this risk. The Rostock cornea module (Heidelberg Engineering) that is an add-on to the HRT II has this risk, but it produces clearer images.

The driving hope is that the measures proposed in this article, or variants of them, will serve as indicators of the development of underlying precursory risk factor conditions, that could be used to evaluate whether a patient ought to be treated. It cannot be concluded that a correlation has been established between the proposed measures and neuropathy progression, because these data are only of limited extent and preliminary. On the other hand, the plots shown in Figures 13 and 14 show a potential and warrant further study.

Based on the wide intersubject variability shown by the vertical extent of the boxes in Figures 13 and 14, it seems daunting to expect any absolute index that has any correlation with risk factor. For such an absolute index, the boxes from one level to the next would have to be nonoverlapping. The idea of establishing a baseline set of measures of a patient and then evaluating a progression from follow-up examinations does seem plausible, albeit not proven.

Work is needed to reduce the required percentage changes shown in Table 3. The smaller that these percentage changes can be made, the more sensitive will be the clinical test. For the type of baseline and follow-up examination indices described in the paragraph above, it is important to improve the technology so that the percentages in column 2 are a fraction of the percentages in columns 3 and 4 of Table 5. Work is also needed to improve the repeatability that is indicated by the *P*-values in Table 2. Higher *P*-values are needed in more of the entries, which would indicate that the means of the measures are virtually unchanged (statistically speaking) from one data collection (i.e., examination) to the next. By "work" we mean that the following factors should be designed, implemented, and tested. Better patterns (e.g., Figs. 5, 6) are needed that provide repeatability measurement means. Variants on measures, besides those in Figures 13 and 14, may provide smaller effect sizes. There are many measures that we thought about after completing this study. One such measure is a fragmentation measure that counts the number of fragmented branch segments after automatically detecting and rejoining branches that

have been segmented due to contrast degradation. Another such measure is a linear combination of several of the existing measures.

Although this work is an early step that offers promise, improvement of the repeatability implied by Table 2 and the sensitivity implied by Table 5 is needed for the clinical use of the algorithm. Tables 2 and 5 are only estimates, so we do not know what the true repeatability and sensitivities may be. Even so, if Table 5 is assumed to be prototypical of how the currently designed measures would perform, the percentages shown in column 2 must be a small fraction of the percentages shown in columns 3 and 4. The purpose of any clinical test, ideally, would be to detect that the underlying neuropathic condition, or risks thereof, are worsening before symptoms worsen, and if the percentages in column 2 are on the same order as the percentages in columns 3 and 4, then the most we could hope for is that the neuropathic condition can only be confirmed. Improved methods, such as linear combinations of the currently designed measures, improved image collection, and improved fiber filtering may provide these needed sensitivity improvements.

References

1. National Institute of Neurologic Disorders and Stroke. Peripheral Neuropathy Fact Sheet. Bethesda, MD: NINDS http://www.ninds.nih.gov/disorders/peripheralneuropathy/detail_peripheralneuropathy.htm/. Accessed July 22, 2010.
2. National Diabetes Statistics 2007, http://diabetes.niddk.nih.gov/dm/pubs/statistics/#allages. Accessed July 22, 2010.
3. Duby JJ, Campbell RK, Setter SM, White JR, Rasmussen KA. Diabetic neuropathy: an intensive review. *Am J Health Syst Pharm.* 2004;61(2):160–176.
4. Samii A, Unger J, Lange W. Vascular endothelial growth factor expression in peripheral nerves and dorsal root ganglia in diabetic neuropathy in rats. *Neurosci Lett.* 1999;262(3):159–162.
5. Schratzberger P, Walter D, Rittig K, et al. Reversal of experimental diabetic neuropathy by VEGF gene transfer. *J Clin Invest.* 2001; 107(9):1083–1092.
6. Bosi E, Conti M. Vermigli, C, et al. Effectiveness of frequency-modulated electromagnetic neural stimulation in the treatment of painful diabetic neuropathy. *Diabetologia.* 2005;48:817–823.
7. Rosenburg ME, Tervo TMT, Immonen IJ, Müller LJ, Grönhagen-Riska C, Vesaluoma MH. Corneal structure and sensitivity in type 1 diabetes mellitus. *Invest Ophthalmol Vis Sci.* 2000;41(10):2915–2921.
8. Kallinikos P, Berhanu M, O'Donnell C, Boulton AJM, Efron N, Malik RA. Corneal nerve tortuosity in diabetic patients with neuropathy. *Invest Ophthalmol Vis Sci.* 2004;45(2):418–422.
9. Guthoff RF, Baudouin C, Stave J. *Atlas of Confocal Laser Scanning In-vivo Microscopy in Ophthalmology.* New York: Springer-Verlag; 2006.
10. Bailey NJ. *Statistical Methods in Biology.* 3rd ed. Cambridge, UK: Cambridge University Press; 1995.
11. Myers JL, Well AD. *Research Design and Statistical Analysis.* 2nd ed. Mahwah, NJ: Lawrence Erlbaum Associates; 2003.
12. Scarpa F, Grisan E, Ruggeri A., Automatic recognition of corneal nerve structures in images from confocal microscopy. *Invest Ophthalmol Vis Sci.* 2008;49(11):4801–4807.
13. Staal J, Abramoff MD, Niemeijer M, Viergever MA, van Ginneken B. Ridge-based vessel segmentation in color images of the retina. *IEEE Trans Med Imaging.* 2004;23:501–509.
14. Giardina CR, Dougherty ER. *Morphological Methods in Image and Signal Processing.* Englewood Cliffs: Prentice Hall; 1988.
15. Grisan E, Foracchia M, Ruggari A. A novel method for the automatic evaluation of retinal vessel tortuosity. *Proceedings of the 25th An-*

nual International Conference of the IEEE-EMBS, Cancun, Mexico; September 17–21, 2003. New York: IEEE; 2003:866–869.

16. Hart W, Goldbaum M, Côté B, Kube P, Nelson M. Automated measurement of retinal vascular tortuosity. *Proceedings of the AMIA* (American Medical Informatics Association) *Fall Conference*. Bethesda, MD: AMIA; 1997:459–463.

17. He W, Hamilton T, Cohen A, Holmes T, Turner J, Roysam B. Automated three-dimensional tracing of HRP stained neurons from a stack of brightfield optical slices. *Microsc Microanal.* 2003;9:296–310.

18. He X, Kischell E, Rioult M, Holmes TJ. Three dimensional thinning algorithm that peels the outmost layer with application to neuron tracing. *J Comp Assist Microsc.* 2000;10(3):123–135.

19. Reed T, Vernon A, Johnson, AM. Sex difference in brain nerve conduction velocity in normal humans. *Neuropsychologia.* 2004;42(12):1709–1714.

20. Licursi de Alcântara AC, Salgado HC, Fazan VPS. Morphology and morphometry of the vagus nerve in male and female spontaneously hypertensive rats. *Brain Res.* 2008;1197:170–180.

21. Muglia U, Vita G, Laura R, Mammola CL, Germana G. Morphometric comparison between controlateral sciatic nerves in the male and female rabbit. *Anat Histol Embryol.* 2007;26(2):147–150.

22. Li B, Qiao G, Feng B, Zhao R, Lu Y, Schild J. Electrophysiological and neuroanatomical evidence of sexual dimorphism in aortic baroreceptor and vagal afferents in rat. *Am J Physiol Regul Integr Comp Physiol.* 2008;295:R1301–R1310.

23. Mowlavi A, Cooney D, Febus L, Khosraviani A, Wilhelmi BJ, Akers G. Increased cutaneous nerve fibers in female specimens. *Plast Reconstr Surg.* 2005;116(5):1407–1410.

24. McLachlan EM. The components of the hypogastric nerve in male and female guinea pigs. *J Auton Nerv Syst.* 1985;13(4):327–342.

25. Reinisch CM, Traxler H, Piringer S, Tangl S, Nader A, Tschachler E. Rarefraction of the peripheral nerve network in diabetic patients is associated with a pronounced reduction of terminal Schwann cells. *Diabetes Care.* 2008;31(6):1219–1221.

26. Auran JD, Koester CJ, Kleinman NJ, et al. Scanning slit confocal microscopic observation of cell morphology and movement within the normal human anterior cornea. *Ophthalmology.* 1995;10:33–41.

27. Patel DV, McGhee CNJ. Mapping of the normal human cornea nerve plexus by in vivo laser scanning confocal microscopy. *Invest Ophthalmol Vis Sci.* 2005;46(12):4485–4488.

28. Patel DV, McGhee CNJ. In vivo laser scanning confocal microscopy confirms that the human corneal sub-basal nerve plexus is a highly dynamic structure. *Invest Ophthalmol Vis Sci.* 2008;49(8):3409–3412.

29. Gabriel JM, Erne B, Pareyson D, Sghirlanzoni A, Taroni F, Steck AJ. Gene dosage effects in hereditary peripheral neuropathy. Expression of peripheral myelin protein 22 in Charcot-Marie-Tooth disease type 1A and hereditary neuropathy liability to pressure palsies nerve biopsies. *Neurology.* 1997;49(6):1636–1640.