The Reliability and Reproducibility of Corneal Confocal Microscopy in Children

Daniele Pacaud,1–3 Kenneth G. Romanchuk,1,2 Mitra Tavakoli,4 Claire Gougeon,1–3 Heidi Virtanen,1–3 Maryam Ferdousi,4 Alberto Nettel-Aguirre,1–3 Jean K. Mah,1–3 Rayaz A. Malik4,5

1Cumming School of Medicine, University of Calgary, Calgary, Alberta, Canada
2Alberta Children’s Hospital, Calgary, Alberta, Canada
3Alberta Children’s Hospital Research Institute, Calgary, Alberta, Canada
4Centre for Endocrinology and Diabetes, Institute of Human Development, University of Manchester, Manchester, United Kingdom
5Weill Cornell Medical College, Al Rayyan, Qatar

Correspondence: Daniele Pacaud, Alberta Children’s Hospital, 2888 Shaganappi Trail NW, Calgary, AB T3B 6A8, Canada; daniele.pacaud@albertahealthservices.ca.
Submitted: March 30, 2015
Accepted: June 20, 2015

PURPOSE. To assess the image and patient level interrater agreement and repeatability within 1 month for corneal nerve fiber length (CNFL) measured using in vivo corneal confocal microscopy (IVCCM) in children.

METHODS. Seventy-one subjects (mean [SD] age 14.3 [2.6] years, range 8–18 years; 44 with type 1 diabetes and 27 controls; 36 males and 35 females) were included. 547 images (~6 images per subject) were analyzed manually by two independent and masked observers. One-month repeat visit images were analyzed by a single masked observer in 21 patients. Automated image analysis was then performed using a specialized computerized software (ACCMetrics).

RESULTS. For CNFL, the ICC (95% CI) were 0.94 (0.93–0.95) for image-level, 0.86 (0.78–0.91) for patient-level, and 0.88 (0.72–0.95) for the 1-month repeat assessment, and the Bland-Altman plots showed minimal bias between observers. Although there was excellent agreement between manual and automated analysis according to an ICC 0.89 (0.82–0.93), the Bland-Altman plot showed a consistent bias with manual measurements providing higher readings.

CONCLUSIONS. In vivo corneal confocal microscopy image analysis shows good reproducibility with excellent intra-individual and inter-individual variability in pediatric subjects. Since the image-level reproducibility is stronger than the patient-level reproducibility, refinement of the method for image selection will likely further increase the robustness of this novel, rapid, and noninvasive approach to detect early neuropathy in children with diabetes. Further studies on the use of IVCCM to identify early subclinical neuropathy in children are indicated.

Keywords: type 1 diabetes, children and adolescents, diabetic neuropathy, corneal confocal microscopy, corneal nerve fiber length

The worldwide burden of diabetes and its complications in children continues to climb due to the rise in the incidence of type 1 and type 2 diabetes.1 According to the Center for Disease Control National Diabetes Statistics Report 2014,2 over 20,000 children and youth less than 20 years of age are diagnosed with diabetes each year in the United States alone, with 78% having type 1 diabetes. Neuropathy is a common complication in adults with diabetes3 leading to neuropathic pain, autonomic dysfunction, foot ulceration, and amputation.1–6 In contrast to adults, children, and adolescents generally do not have signs or symptoms of diabetic neuropathy.7–9 While overt diabetic neuropathy is rarely present in children with type 1 diabetes, subclinical diabetic neuropathy has been estimated to occur in up to 68% of all children with type 1 diabetes with a duration of 5 years or longer, and up to 25% of pediatric patients with newly diagnosed diabetes have abnormal nerve conduction studies.10

National and international clinical practice guidelines for children with type 1 diabetes11–13 recommend yearly screening for diabetic neuropathy at puberty and after 5 years of diabetes duration. According to prevalence statistics and recommendations, Hirschfeld et al.13 estimated that several hundred children with diabetes are screened for the presence of peripheral diabetic neuropathy every day in the United States. However, traditional tools such as the use of Semmes-Weinstein monofilament to screen for peripheral neuropathy are often found to be unreliable, especially in children.14,15 For example, in a recent study of 88 children with type 1 diabetes, the diagnostic use and interrater agreement were very low for both the monofilament and the tuning fork.16 Assessment of diabetic neuropathy in children, therefore, remains a challenge due to the absence of clinical symptoms, lack of good pediatric normative data, difficulties with time-consuming, technically more challenging, and sometimes uncomfortable electrophysiologic tests such as nerve conduction studies (NCS), and lack of reliability and diagnostic sensitivity of screening techniques deployed in adults.17 In vivo corneal confocal microscopy (IVCCM) is increasingly recognized as an ideal noninvasive surrogate marker of diabetic neuropathy.
neuropathy in adults, but its use in children has not been widely studied. Of the different parameters that can be measured with IVCCM, corneal nerve fiber length (CNFL) seems to be the most reliable to date. In our recent experience it has been shown to have use in longitudinal studies to predict the development of neuropathy in adults with type 1 diabetes. Furthermore, we have also shown that it may be particularly useful in detecting autonomic dysfunction and particularly overt autonomic neuropathy.

However, before IVCCM can be proposed as a diagnostic test in children with diabetes, it is important to establish the reliability and reproducibility of this technique in this age group. Because of the unique physiologic and developmental issues in the pediatric population, any new diagnostic test should be specifically validated in children before its use in both research and clinical practice. In the case of IVCCM, although adult studies have provided some information on expected changes with age leading to slow decline in number of fibers, it is not presently known at what age or developmental stage that this decline occurs. Further, before having studied IVCCM in children, the assumption that image analysis will be identical as in adults can potentially lead to erroneous interpretation. Therefore, the main objective of this analysis was to assess the interobserver agreement and repeatability of CNFL measurements in IVCCM images from children.

**Methods**

**Study Subjects**

This study was approved by the Conjoint Health Research Ethics Board at the University of Calgary (Calgary, Alberta, Canada) and the research was conducted in accordance with Declaration of Helsinki. As part of a larger study on use of IVCCM in children with diabetes, children with a history of at least 5 years of type 1 diabetes aged 8 to 18 years and healthy controls of the same age were invited to participate. Subjects were recruited through the Alberta Children’s Hospital Diabetes Clinic. Healthy control subjects were recruited through advertisement posters set in general pediatric or pediatric ophthalmology clinics and through word of mouth from siblings and friends of children with diabetes. After informed consent was obtained, basic demographic and diabetes related data were collected through questionnaire and chart review. As use of contact lenses is not associated with changes in corneal nerve morphology, it was not part of the exclusion criteria for our study. Each participant completed an assessment of neuropathy including clinical symptoms and signs, nerve conduction studies, quantitative sensory and autonomic function tests. Each subject also underwent an assessment of IVCCM; a subgroup of 21 subjects are shown in Table 1. There were no significant differences between groups for age, sex, ethnicity, tobacco smoking, or use of contact lenses.

To assess the interrater reliability (Table 2), we compared CNFL results for agreement between the two observers. For this purpose, 547 images (303 images from individuals with diabetes and 244 images from healthy subjects) were analyzed. For image-level results, there was excellent reliability between observers with an ICC of 0.94 (0.93–0.95). Next we compared the average CNFL results for each patient (patient-level; from 71 subjects: 44 with diabetes and 27 healthy subjects) between the two observers and it demonstrated excellent reliability with an ICC of 0.86 (0.78–0.91). Intraclass correlation coefficient (ICC) as a measure of repeatability and Bland and Altman plots were computed to illustrate the agreement between measurement methods or observers. Intraclass correlation coefficient were considered excellent if they ranged between 0.8 and 1, and very good between 0.6 and 0.79.

**Image Analysis**

Images were manually analyzed using a custom designed nerve analysis software package (CCMetrics V2, MA Dabbah, Imaging Science; University of Manchester, Manchester, UK). For each image, all nerve fibers and branches visible within the frame of an image were traced on a tablet computer using a digital pen and the nerve fiber length (CNFL) were quantified. Hence, CNFL is defined as the total length of all nerve fiber and branches per square millimeter of corneal tissue (mm/mm²).

Fully automated analysis of each image was also performed using the same software (CCMetrics V2, MA Dabbah, Imaging Science). Images were selected by one examiner (MT) and analyzed by a single individual in Manchester (MF) while a single individual selected and analyzed images in Calgary (KR). All were blinded to the status of the subjects.

**Statistical Analysis**

All analyses were done using IBM SPSS Statistics (Version 19; IBM Corp., Armonk, NY, USA) while the Bland and Altman plots were computed using the R project for statistical computing. Descriptive statistics are presented as mean and SD for numerical/continuous variables and percentages for categorical variables. For all subjects, image-level (for those images analyzed by both raters) and patient-level (average of images selected by the observer for a single patient) results were compared between both observers. For patient-level results, because there was no significant difference between the results of the left and right eye, results from both eyes were averaged together (one value of the average nerve fiber length per patient). Comparisons between the two observers and two examinations of the same patient as well as comparison between manual analysis and automated analysis were assessed with the intraclass correlation coefficient (ICC) as a measure of repeatability and Bland and Altman plots were computed to illustrate the agreement between measurement methods or observers. Intraclass correlation coefficient were considered excellent if they ranged between 0.8 and 1, and very good between 0.6 and 0.79.

**Results**

Seventy-one children (51% males; mean age 14.3 [SD 2.6] years) were included in this study. Forty-four had type 1 diabetes with average duration of 9.2 (SD 2.7) years, and 27 were healthy control subjects. Additional characteristics of the sample are shown in Table 1. There were no significant differences between groups for age, sex, ethnicity, tobacco smoking, or use of contact lenses.
The CNFL magnitudes are higher (Fig. 1). The overall spread of the CNFL values was similar in the image-level and patient-level CNFL measurements when comparing the two observers.

When comparing manual with automated analysis, CNFL measurements at the subject-level on the first visit were 22.0 (5.5) mm/mm² for manual analysis versus 15.1 (4.3) mm/mm² for automated analysis. The resulting ICC was 0.89 (0.82–0.93) showing excellent reliability despite manual measurements providing consistently higher values. This is also illustrated in the Bland-Altman plot (Fig. 2) showing a bias of −6.7 (−11.9 to −1.5) between manual and automated analysis. On the 1-month repeated scans, the automated analysis had similar ICC as the manual analysis (0.87 [0.68–0.95] vs. 0.89 [0.72–0.95]), while the bias on the Bland-Altman plots was minimal for both (automated −0.2 and manual −0.1; Plots not shown).

**DISCUSSION AND CONCLUSIONS**

Three previous groups have reported on the reliability and reproducibility of IVCCM in adults with diabetic neuropathy. To our knowledge, this study is the largest study to date on IVCCM reliability for the measurement of CNFL and the first report of IVCCM reliability in children. As IVCCM is a promising rapid and noninvasive tool to detect diabetic neuropathy in adults, it is important to assess its use in children, as they may show the earliest deficits.

The criteria for diagnosing diabetic neuropathy requires the presence of symptoms and signs of peripheral neuropathy, or symptoms or signs with abnormal testing from NCS, quantitative sensory testing (QST) or autonomic testing in individuals with diabetes after excluding other causes of neuropathy. Subclinical neuropathy is defined by abnormal testing only. This definition based on clinical testing requires multiple tests, which are time consuming, more difficult to apply to children and not always easily accessible in pediatric clinical settings.

**Table 1. Subject Characteristics**

<table>
<thead>
<tr>
<th></th>
<th>T1DM</th>
<th>Control</th>
<th>T1DM + Controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. participants</td>
<td>44</td>
<td>27</td>
<td>71</td>
</tr>
<tr>
<td>Age, y</td>
<td>14.8 (2.3)</td>
<td>13.6 (2.9)</td>
<td>14.3 (2.6)</td>
</tr>
<tr>
<td>Sex, M/F (%/%)</td>
<td>21/25 (47.7/52.3)</td>
<td>15/12 (55.6/44.4)</td>
<td>36/35 (50.7/49.5)</td>
</tr>
<tr>
<td>Smoking, yes/no (%/%)</td>
<td>1/45 (2.3/97.7)</td>
<td>1/26 (3.7/96.5)</td>
<td>2/69 (2.8/97.2)</td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>40 (90.9)</td>
<td>23 (85.2)</td>
<td>63 (88.7)</td>
</tr>
<tr>
<td>Asian</td>
<td>5 (6.8)</td>
<td>3 (11.1)</td>
<td>6 (8.5)</td>
</tr>
<tr>
<td>Black</td>
<td>1 (2.3)</td>
<td>0 (0.0)</td>
<td>1 (1.4)</td>
</tr>
<tr>
<td>Other</td>
<td>0 (0.0)</td>
<td>1 (3.7)</td>
<td>1 (1.4)</td>
</tr>
<tr>
<td>Contact lenses, yes/no (%/%)</td>
<td>6/38 (13.6/86.4)</td>
<td>1/26 (3.7/96.3)</td>
<td>7/64 (9.9/90.1)</td>
</tr>
<tr>
<td>Duration of diabetes, y</td>
<td>9.2 (2.7)</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Age at diagnosis, y</td>
<td>5.6 (3.0)</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>A1C (%)</td>
<td>8.9 (1.7)</td>
<td>N/A</td>
<td>N/A</td>
</tr>
</tbody>
</table>

Data presented as mean (SD) or number (percentages).
* Upper limit of normal for individuals without diabetes is 6.1%.

**Table 2. Comparison of CNFL Measurements Between Observer 1 and Observer 2**

<table>
<thead>
<tr>
<th></th>
<th>CNFL Observer 1</th>
<th>CNFL Observer 2</th>
<th>ICC (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Image level</td>
<td>23.9 (6.2)</td>
<td>24.9 (6.6)</td>
<td>0.94 (0.93–0.95)</td>
</tr>
<tr>
<td>Patient level</td>
<td>22.3 (4.8)</td>
<td>24.9 (5.4)</td>
<td>0.86 (0.78–0.91)</td>
</tr>
</tbody>
</table>

Corneal nerve fiber length are presented mean (SD) in millimeters per millimeters squared.

**Figure 1.** Bland-Altman plots for CNFL (mm/mm²) indicating the level of agreement between two observers for image level (A) and for patient level (B) measurements.
Further, many of the screening tests used in adults have not been well validated in children and do not have age appropriate norms. For example, although the use of the 10-g Semmes–Weinstein microfilament has been suggested as a good screening tool for diabetic neuropathy and a predictor for risk of amputation in adults, it has been found to be invalid as a screening tool in children. Chuback et al. have shown that vibration and tactile perception of youth with type 2 diabetes failed to identify those at risk of amputation in young adulthood, supporting the need for a reliable screening tool to detect subclinical or early neuropathy in children. Hence, there is a need for pediatric specific norms to be established. The single, stand-alone test considered as the gold standard to assess diabetic neuropathy is a skin biopsy to demonstrate the presence and evolution of diabetic neuropathy. As such, similar to our experience in type 1 children and adolescents, Sellers et al. also found that IVCCM was quick, easy, and well accepted by a group of teenagers with type 2 diabetes. Most studies using IVCCM report on four different nerve morphology parameters: nerve fiber density, nerve fiber branch density, nerve fiber length, and nerve fiber tortuosity. Of these four parameters, CNFL has been shown to have the best reproducibility and the best relationship to the severity of diabetic neuropathy. As in adults, we now demonstrate excellent interobserver agreement for both single images and average patient-level CNFL measurements. However, the patient-level ICC was lower and bias was larger (MT trained KR for image selection and analysis on a different set of images not included in this study). Thereafter, although the protocols for image-level analysis was done similarly between the two observers, image selection differed with one observer selecting five to eight best images, while the other observer always aimed for eight images per eye. This is the likely explanation for the lower agreement as shown both by lower ICC and the bias on the Bland-Altman plots for patient level measurements. That is, the averages used in the patient-level analysis were not necessarily the averages of only the images used for the image level analysis. Further refinement for image selection criteria and use of automated analysis will likely further increase the robustness of CCM measurements. However, the cutoff for normal values will be lower when images are analyzed by an automated program. This reinforces the need to establish good normative data specific to both age groups and the method of analysis (automated versus manual).

Repeated image analysis also shows good reliability. The relatively lower ICC can be explained by the known migration of nerve fibers in the corneal subbasal layer and the fact that the IVCCM cannot recapture exactly the same central area. On the 1-month repeat, our ICC were higher than those reported by Hertz et al. This may be explained by less variability of CNFL in our patients as most had minimal neuropathy, whereas their sample included healthy adult subjects and subjects with the whole spectrum of severity of diabetic neuropathy.

Limitations of the current study include differences in image collection and sampling techniques. The slight variation in image selection between the two observers may be seen as a limitation of our study; however, we believe this more accurately reflects what may happen in the ‘real world’ when different operators and different centers undertake IVCCM. As per Hertz et al., we have chosen to present the patient-level data as the average of both left and right eye, which is in contradiction with optometry conventions, where unilateral measurements are provided for each eye. However, studies to date have not found significant differences between the left and right eyes. We have adopted the same protocols for image acquisition and analysis as the majority of recently published studies deploying IVCCM, including a large normative study.

In conclusion, we show that IVCCM is a reproducible and reliable technique with excellent intraindividual and interindividual variability in pediatric subjects. If proven to be equal or superior to nerve conduction studies to detect diabetic neuropathy, this novel, rapid, and noninvasive approach could therefore be deployed in screening and assessing the effect of therapies in children with diabetic neuropathy.

**Acknowledgments**

The authors thank the families and patients who agreed to participate in the study.

Supported by a grant from Juvenile Diabetes Research Foundation International (Grant: 17-2008-1032; New York, NY, USA).

Disclosure: D. Pacaud, None; K.G. Romanchuk, None; M. Tavakoli, None; C. Gougeon, None; H. Virtanen, None; M. Ferdousi, None; A. Nettel-Aguirre, None; J.K. Mah, None; R.A. Malik, None.

**References**


Reliability of CCM in Children


