Journal of Diabetes and Its Complications 31 (2017) 1066–1073

Contents lists available at ScienceDirect

Journal of Diabetes and Its Complications

journal homepage: WWW.JDCJOURNAL.COM

Agreement between automated and manual quantification of corneal nerve fiber length: Implications for diabetic neuropathy research

Daniel Scarr [a], Leif E. Lovblom [a], Ilia Ostrovski [a], Dylan Kelly [b], Tong Wu [a], Mohammed A. Farooqi [a], Elise M. Halpern [a], Mylan Ngo [c], Eduardo Ng [c], Andrej Orszag [a], Vera Bril [c], Bruce A. Perkins [a,b,*]

[a] Lunenfeld-Tanenbaum Research Institute, Mount Sinai Hospital, Toronto, Ontario, Canada
[b] Division of Endocrinology and Metabolism, Department of Medicine, Mount Sinai Hospital, University of Toronto, Toronto, Ontario, Canada
[c] The Ellen and Martin Prosserman Centre for Neuromuscular Diseases, Krembil Neuroscience Centre, Division of Neurology, Department of Medicine, University Health Network, University of Toronto, Toronto, Ontario, Canada

ARTICLE INFO

Article history:
Received 29 March 2016
Received in revised form 24 June 2016
Accepted 22 July 2016
Available online 26 July 2016

Keywords:
Corneal confocal microscopy
Corneal nerve fiber length
Diabetes
Neuropathy
Diabetic neuropathy
Biomarkers

ABSTRACT

Aims: Quantification of corneal nerve fiber length (CNFL) by in vivo corneal confocal microscopy represents a promising diabetic neuropathy biomarker, but applicability is limited by resource-intensive image analysis. We aimed to evaluate, in cross-sectional analysis of non-diabetic controls and patients with type 1 and type 2 diabetes with and without neuropathy, the agreement between manual and automated analysis protocols.
Methods: Sixty-eight controls, 139 type 1 diabetes, and 249 type 2 diabetes participants underwent CNFL measurement (N = 456). Neuropathy status was determined by clinical and electrophysiological criteria. CNFL was determined by manual ($CNFL_{Manual}$, reference standard) and automated ($CNFL_{Auto}$) protocols, and results were compared for correlation and agreement using Spearman coefficients and the method of Bland and Altman ($CNFL_{Manual}$ subtracted from $CNFL_{Auto}$).
Results: Participants demonstrated broad variability in clinical characteristics associated with neuropathy. The mean age, diabetes duration, and HbA1c were 53 ± 18 years, 15.9 ± 12.6 years, and 7.4 ± 1.7%, respectively, and 218 (56%) individuals with diabetes had neuropathy. Mean $CNFL_{Manual}$ was 15.1 ± 4.9 mm/mm^2, and mean $CNFL_{Auto}$ was 10.5 ± 3.7 mm/mm^2 ($CNFL_{Auto}$ underestimation bias, −4.6 ± 2.6 mm/mm^2 corresponding to −29 ± 17%). Percent bias was similar across non-diabetic controls (−33 ± 12%), type 1 (−30 ± 20%), and type 2 diabetes (−28 ± 16%) subgroups (ANOVA, $p = 0.068$), and similarly in diabetes participants with and without neuropathy. Levels of $CNFL_{Auto}$ and $CNFL_{Manual}$ were both inversely associated with neuropathy status.
Conclusions: Although $CNFL_{Auto}$ substantially underestimated $CNFL_{Manual}$, its bias was non-differential between diverse patient groups and its relationship with neuropathy status was preserved. Determination of diagnostic thresholds specific to $CNFL_{Auto}$ should be pursued in diagnostic studies of diabetic neuropathy.

1. Introduction

Diabetic sensorimotor polyneuropathy (DSP) involves progressive, diffuse and length-dependent injury to peripheral nerves and affects up to 50% of people with diabetes (Boulton, 2005). Its underlying causes are complex and include, but are not limited to, chronic hyperglycemia (Boulton et al., 2005). The progression of DSP is associated with significant morbidity and costs (Gordois, Scuffham, Shearer, Oglesby, & Tobian, 2003). It can cause pain, imbalance, and foot deformity; in later stages it can result in infection, ulceration, and amputation. In view of a long subclinical phase there is an urgent need for a biomarker of early DSP for use in clinical practice and implementation in future trials of therapies aimed at preventing DSP onset and progression (Canadian Diabetes Association Clinical Practice Guidelines Expert Committee et al., 2013; Ziegler & Luft, 2002).

From a research perspective and frequently in clinical practice, the diagnosis of DSP requires confirmation by electrodiagnostic nerve conduction studies (NCS) (England et al., 2005; Tesfaye et al., 2010). However, NCS may have limitations in detecting early, pre-clinical stages in which abnormalities in small nerve fiber function or structure are more prominent than those in large nerve fibers (Breiner, Lovblom, Perkins, & Bril, 2014; Smith & Singleton, 2008;

Disclosure statement: B.A.P. has received speaker honoraria from Medtronic Inc., Johnson and Johnson, Roche, GlaxoSmithKline Canada, Novo Nordisk and Sanofi; has received research grant support from Medtronic and Boehringer Ingelheim. The remaining authors have no relevant conflicts of interest to disclose.

* Corresponding author at: Endocrinology and Metabolism, Lunenfeld-Tanenbaum Research Institute, University of Toronto, Leadership Sinai Centre for Diabetes, Mount Sinai Hospital, L5209-60 Murray Street Box 16, Toronto, ON, Canada M5T3L9. Tel.: +1 416 586 8763; fax: +1 647 826 1528.
E-mail address: bperkins@mtsinai.on.ca (B.A. Perkins).

http://dx.doi.org/10.1016/j.jdiacomp.2016.07.024
1056-8727/

Sumner, Sheth, Griffin, Cornblath, & Polydefkis, 2003). The current reference standard for measuring small fiber morphological abnormality is skin punch biopsy, which is capable of detecting morphological changes of intra-epidermal nerve fibers (England et al., 2009). Though administered well in research and in some clinical settings, a skin biopsy can be considered invasive for DSP screening and may not be generalizable to a screening program for all patients with diabetes. As an alternative, in vivo Corneal Confocal Microscopy (IVCCM) has emerged as a non-invasive means of examining small nerve fiber morphology (Oliveira-Soto & Efron, 2001; Quadrado, Popper, Morgado, Murta, & Van Best, 2006; Quattrini et al., 2007; Rozsa & Beuerman, 1982; Tavakoli et al., 2010). This technique allows for the visualization of small nerve fibers located in the sub-basal nerve plexus of the cornea, anterior to the Bowman's layer. The morphology of these nerve fibers is believed to closely reflect that of the nerves affected by DSP (Quattrini et al., 2007). The morphologic parameter that has shown particular promise as a potential biomarker for early-stage DSP is corneal nerve fiber length (CNFL). CNFL is the most reproducible IVCCM parameter, most strongly associated with the presence of DSP, and may be effective at predicting future onset of DSP (Ahmed et al., 2012; Efron et al., 2010; Hertz et al., 2011; Lovblom et al., 2015; Pritchard et al., 2015).

Although simple protocols for image acquisition have been successfully implemented in research settings, the key barrier to the generalizability of IVCCM to clinical practice and research is its resource-intensive image analysis procedure. Current gold-standard methodology is extremely time consuming and requires expertise of a trained technician, who must first identify nerves on the IVCCM images and manually trace them using a graphic pen and tablet before the morphological parameters can be quantified. In light of this, novel and fully-automated software has been developed that eliminates the need for this manual tracing process (Dabbah, Graham, Petropoulos, Tavakoli, & Malik, 2011). Such an automated approach is highly reproducible (Ostrovski et al., 2015; Pacaud et al., 2015). However, previous studies of various patient populations with and without diabetes, including one in adolescents (Pacaud et al., 2015), have found varying degrees of underestimation of the manual protocol by the new automated protocol (Chen et al., 2015; Dabbah et al., 2011; Dehghani et al., 2014; Ostrovski et al., 2015; Parissi et al., 2013; Petropoulos et al., 2014). This variation may have arisen from differences in the measurement protocols, or from differences in sample size. Additionally, the previous studies did not primarily aim to quantitatively assess the agreement between the two protocols or to fully evaluate the nature of intrinsic measurement bias (Kottner et al., 2011). Therefore, the purposes of this study were to verify the agreement between automated and manual protocols by evaluating the automated software in a large and diverse study population consisting of 456 participants, and to examine if the measurement bias was differential according to presence or absence of diabetes, diabetes type, and the presence or absence of neuropathy. In using a generalizable approach to the automated IVCCM protocol intended for broad clinical use, we aimed to examine its correlation and agreement with the current resource-intensive reference manual protocol.

2. Materials and methods

2.1. Study population

The cross-sectional baseline data of 456 participants from two cohorts were examined in this analysis. The first cohort consisted of 139 type 1 diabetes (T1D) participants and 68 age- and gender-matched non-diabetic controls from the Toronto Longitudinal Neuropathy Cohort, funded by the JDRF (Operating Grant 17-2008-715) conducted between November 2008 and July 2013 (Ahmed et al., 2012; Halpern et al., 2013; Lovblom et al., 2015). The second cohort consisted of 249 type 2 diabetes (T2D) participants recruited for a study funded by the Canadian Diabetes Association (Operating Grant OG-3-10-3123-BP) conducted between November 2010 and May 2013 (Farooqi et al., 2016). Participants with diabetes in both cohorts were recruited consecutively from the Endocrinology and Metabolism Clinic and the Prosserman Family Neuromuscular Clinic at the Toronto General Hospital (University Health Network, Toronto, Ontario, Canada), and non-diabetic controls were recruited through family members and friends of the diabetes participants, and through community advertisement. Both studies investigated the relationship between IVCCM and DSP, with accrual strategies aiming to include participants with a broad spectrum of nerve injury ranging from no detectable nerve injury to severe DSP. This was accomplished using a stratified accrual strategy according to the Toronto Clinical Neuropathy Score (TCNS), a validated grading system to evaluate history and physical exam components that permitted tracking of the number of subjects likely to have absent, mild, moderate, and severe neuropathy at the time of study accrual (Bril & Perkins, 2002). Potential participants were excluded if they had history of non-diabetic neuropathy, eye infection or other conditions that precluded safe IVCCM examination, or if they had an allergy to the ocular anesthetic used during the IVCCM exam. Study participants were ≥18 years of age and provided written informed consent. The protocol and consent procedures for both studies were approved by the research ethics board of the Toronto General Hospital Research Institute.

2.2. IVCCM examination

Participants underwent bilateral examination of the sub-basal plexus anterior to the Bowman's layer of the cornea using the Rostock Cornea Module of the Heidelberg Tomograph III (Heidelberg Engineering, Smithfield, RI, USA) (Tavakoli & Malik, 2011). The methods of this examination have been described previously (Ahmed et al., 2012; Lovblom et al., 2015; Ostrovski et al., 2015). Compared to the image acquisition procedure adopted by others (Edwards et al., 2012; Petropoulos et al., 2014), we implemented the less operator-dependent 'volume scan' method in which the examiner focuses the microscope on the participant's central cornea and captures images through an automated process that incrementally captures 40 images semi-randomly at increasing depth throughout the thickness of the cornea. Each digital image had dimensions 384 × 384 pixels. All participants underwent IVCCM using the 300-μm^2 field of view (FOV) lens (N = 456), while a subgroup of each cohort (15 non-diabetic controls, 37 type 1 diabetes participants, and 224 type 2 diabetes participants) underwent IVCCM using both a 300-μm^2 and 400-μm^2 FOV lenses, on the same day of the exam (Hume et al., 2012).

2.3. Manual image selection and CNFL quantification

The manual protocol of image analysis required the examiner to select two images per participant, one image from each of the right and left eyes. The criterion for selection was very concise, such that the examiner had to select high-contrast images (free of artifacts such as corneal folding and stromal dendrite cells) deemed to yield the highest CNFL (Ahmed et al., 2012; Halpern et al., 2013). Although there is no established standard for how many images to analyze in the IVCCM protocol for DSP, other investigators have analyzed the average $CNFL_{Manual}$ of 4–8 images per participant (Messmer, Schmid-Tannwald, Zapp, & Kampik, 2010; Quattrini et al., 2007; Tavakoli et al., 2010). Due to the amount of time taken to manually trace each image, our protocol aimed to test a variation from the protocol in which only 2 images were selected (Ahmed et al., 2012; Edwards et al., 2012; Tavakoli et al., 2010).

The examiner carrying out the manual protocol used CCMetrics Image Analysis Tool v1.1 (provided by Drs. R Malik and M Dabbah, University of Manchester) to manually trace the nerve fibers and branches in each image. To determine the $CNFL_{Manual}$ represented by each image, an output of the number of pixels occupied by the nerve