Journal of Microscopy

Journal of Microscopy, Vol. 245, Pt 1 2012, pp. 82–89
Received 30 December 2010; accepted 5 August 2011

doi: 10.1111/j.1365-2818.2011.03547.x

Collagen bundle morphometry in skin and scar tissue: a novel distance mapping method provides superior measurements compared to Fourier analysis

PAULINE D.H.M. VERHAEGEN*,†,‡,§, JAN VAN MARLE¶, ANDREAS KUEHNE#,**, HENNIE J. SCHOUTEN‡,††, EAMONN A. GAFFNEY#, PHILIP K. MAINI#,‡‡, ESTHER MIDDELKOOP*,†,§§ & PAUL P.M. VAN ZUIJLEN*,†,‡,§,§§

**Association of Dutch Burn Centres, Beverwijk, The Netherlands*

†Department of Plastic, Reconstructive and Hand Surgery, Red Cross Hospital, Beverwijk, The Netherlands

‡Burn Centre, Red Cross Hospital, Beverwijk, The Netherlands

§Department of Plastic, Reconstructive and Hand Surgery, Academic Medical Centre, Amsterdam, The Netherlands

¶Department of Cell Biology and Histology, Academic Medical Centre, Amsterdam, The Netherlands

#Centre for Mathematical Biology, Mathematical Institute, University of Oxford, Oxford, U.K.

***Institute of Pharmacy and Molecular Biotechnology, University of Heidelberg, Heidelberg, Germany*

††Department of Physiotherapy, Red Cross Hospital, Beverwijk, The Netherlands

‡‡Department of Biochemistry, Oxford Centre for Integrative Systems Biology, University of Oxford, Oxford, U.K.

§§Department of Plastic, Reconstructive and Hand Surgery, VU Medical Centre, Amsterdam, The Netherlands

Key words. Bundle spacing, bundle thickness, collagen, fibrosis, morphometry, scar, skin.

Summary

Histopathological evaluations of fibrotic processes require the characterization of collagen morphology in terms of geometrical features such as bundle orientation thickness and spacing. However, there are currently no reliable and valid techniques of measuring bundle thickness and spacing. Hence, two objective methods quantifying the collagen bundle thickness and spacing were tested for their reliability and validity: Fourier first-order maximum analysis and Distance Mapping, with the latter constituting a newly developed morphometric technique. Histological slides were constructed and imaged from 50 scar and 50 healthy human skin biopsies and subsequently analyzed by two observers to determine the interobserver reliability via the intraclass correlation coefficient. An intraclass correlation coefficient larger than 0.7 is considered as representing good reliability. The interobserver reliability for the Fourier first-order maximum and for the Distance Mapping algorithms, respectively, showed an intraclass correlation coefficient above 0.72 and 0.89. Additionally, we performed an assessment of validity in the form of responsiveness, in particular, demonstrating medium to excellent results via a calculation of the effect size, highlighting that both methods are sensitive enough to measure a treatment effect in clinical practice. In summary, two reliable and valid measurement methods were demonstrated for collagen bundle morphometry for the first time. Due to its superior reliability and more useful measures (bundle thickness and bundle spacing), Distance Mapping emerges as the preferred and more practical method. Nevertheless, in the future, both methods can be used for reliable and valid collagen morphometry of skin and scars, whereas further applications evaluating the quantitative microscopy of other fibrotic processes are anticipated.

Correspondence to: P. P. M. van Zuijlen, Department of Plastic, Reconstructive and Hand Surgery, Red Cross Hospital Beverwijk, Vondellaan 13, 1942 LE Beverwijk. Tel: 0031 251265785; fax: 0031 251265342; e-mail: paul.van.zuijlen@planet.nl

Introduction

The repair of collagen structure plays a central role in wound healing, pathological fibrotic disorders and scar formation. Hence, the histopathological analysis of collagen structure is clinically relevant for diagnosing dermatological diseases of connective tissue (de Vries *et al.*, 2000), distinguishing between different types of scars (Verhaegen *et al.*, 2009) and evaluating fibrotic processes of other viscera, such as heart, lung, kidney and liver (Whittaker *et al.*, 1991; Whittaker *et al.*, 1994). In particular, for dermatological diseases such as scleroderma and lichen sclerosis the collagen bundle orientation and morphology aids in the discrimination between healthy skin and sclerotic dermal tissue. The collagen bundles in sclerotic dermal tissue show an orientation parallel to the epidermis, whereas collagen bundles in healthy skin are considered to be organized in a three-dimensional basket-weave pattern (de Vries *et al.*, 2000). The histological analysis of collagen bundle orientation and morphology in scars provides further insights into the distinction between hypertrophic scars and keloids. We showed that the collagen bundles in keloids are significantly thicker compared to hypertrophic scars (Verhaegen *et al.*, 2009), which is relevant because hypertrophic scars and keloids require distinct therapeutic strategies and are often difficult to discriminate (Niessen *et al.* 1999; Bloemen *et al.*, 2009).

The most common criteria for the analysis of collagen structure are the orientation and the thickness of the collagen bundles. In previous literature, Baak stressed the importance of reproducible and accurate measurement methods for pathological analyses (Baak, 2002). Nowadays, performing reproducible and accurate measurements has become a key focus in pathological practice as well as in clinical and experimental research. In this paper, we will focus on both the reliability and the validity (in the form of responsiveness) of two objective measurement methods for the analysis of collagen bundle thickness and collagen bundle spacing. Reliability refers to 'whether repeated measurements provide similar results?' (Streiner & Norman, 2008a) whereas responsiveness considers 'if a measurement technique is able to identify a meaningful or clinically important change (sensitivity to change)?' (Streiner & Norman, 2008b). Assessment of the responsiveness provides an answer to the question 'whether a measurement method is sensitive enough to be able to measure a treatment effect in clinical practice' (Kazis *et al.*, 1989).

For quantitative evaluation of the collagen bundle orientation, two measurement methods have been presented in literature. Noorlander *et al.* (2002) developed a method for the quantification of individual bundles in the dermis on picrosirius red-stained sections, using the length of the collagen bundles in the plane of the section, with the length of the 10 longest collagen bundles characterizing the orientation of the bundles for example. This method was later applied in an experimental study into the effects of stretch on the collagen orientation of porcine skin (Melis *et al.*, 2002). Collagen bundle orientation was also quantified by a Fourier zeroth-order maximum analysis, which calculates information about spatial organization in images by generating a power spectrum of all the structures present (de Vries *et al.*, 2000). The collagen orientation index can be calculated using the width–length ratio of this power spectrum, that is:

$$\left[1 - \frac{\text{width of the zeroth order power spectrum}}{\text{length of the zeroth order power spectrum}}\right].$$

Zero then corresponds to perfectly random-orientated collagen bundles and one corresponds to perfectly parallel-orientated collagen bundles. Van Zuijlen *et al.* (2002a,b) showed that the zeroth-order maximum of Fourier analysis can be used reliably for determining the orientation of the collagen bundles and that it is a valid measurement method for this purpose. Fourier analysis has been widely applied to determine the orientation of structures and, besides using it for the objective assessment of collagen orientation in diseased skin and scars (van Zuijlen *et al.*, 2002a,b; Har-Shai *et al.*, 2003; van Zuijlen *et al.*, 2003; Ng *et al.*, 2005; Har-Shai *et al.*, 2006; Singer & McClain., 2006,), it has been used to quantify the orientation of structures in other types of tissue, such as ligaments (Chaudhuri *et al.*, 1987), the annulus fibrosus (Guerin & Elliott, 2006) and vascular endothelial cells (Palmer & Bizios, 1997).

The morphometry of collagen bundle thickness and spacing has been investigated by de Vries *et al.* (2000) and Ferdman & Yannas (1993), with the latter study estimating the average collagen bundle diameter by the scattering of laser light through histological sections (Ferdman & Yannas 1993), albeit with a self-confessed simplistic model and no test of reliability. De Vries *et al.* (2000) described a second application of Fourier analysis via the first-order maximum. Measurements using this method were nevertheless partially observer-dependent because manual thresholding was necessary. The Fourier first-order maximum allowed the estimation of 'bundle packing' [in this paper, referred to as the *Bundle Centreline Distance* (BCD_{FFT})], which is the average centre-to-centre distance between the collagen bundles, and this has been applied in clinical studies (van Zuijlen *et al.*, 2002a,b; van Zuijlen *et al.*, 2003). Because, up to now, this measurement method was not tested for its reliability and validity, in this study the reliability and responsiveness (as part of the validation process) of the Fourier first-order maximum analysis was assessed. Unfortunately, the Fourier first-order maximum is only capable of measuring the distance between the centres of the collagen bundles, whereas an objective evaluation of the thickness and spacing of the bundles would provide more practical information. We therefore developed a new (semi-)automated method for measuring collagen bundle thickness and collagen bundle

spacing: Distance Mapping. This method only requires the (optional) selection of the measurement area without additional observer intervention. In Distance Mapping, a distance map is generated after segmentation of an image. Subsequently, the grey values of the pixels in this distance map are measured with a skeleton mask. In summary, the objective of this study is, for the first time, to assess the reliability and validity of measurement methods for collagen bundle morphology. In particular, we explored measurement of the *bundle thickness* and *bundle spacing* of the newly developed method of Distance Mapping and the *Bundle Centreline Distance of the* Fourier first-order maximum algorithm in detail.

Materials and methods

Tissue specimen

The reliability of Distance Mapping and Fourier first-order maximum analysis was tested on 50 consecutive sections from healthy skin and 50 sections from scar tissue. The biopsies consisted of residual patient material from healthy and scarred skin that was collected during plastic surgical corrections, such as breast reductions, abdominoplasties and scar excisions. Scar tissue that was used consisted of normotrophic and hypertrophic scars. The protocols set by the Federation of Dutch Medical Scientific Societies, which are adapted by the coordinating ethics committee in The Netherlands, permit the use of anonymized residual tissue. This study was conducted in accordance with the Helsinki Declaration guidelines. The mean age of patients was 40.9 (SD 12.6) years for the healthy skin biopsies and 30.0 (17.7) years for the scar biopsies.

After harvesting, the biopsies were fixed in 4% formalin for at least 24 h, dehydrated by standard histological procedures and embedded in paraffin. Sections of 5 μm were cut perpendicular to the skin surface, mounted on glass slides and stained with haematoxylin and eosin.

Confocal laser scanning microscopy

We used the fluorescent properties of eosin to avoid dominance of the haematoxylin stained nuclei in a classic bright field image (http://home.earthlink.net/~fluorescentdyes). The slides were imaged with a Leica SP2-AOBS confocal microscope (Leica-Microsystems, Mannheim, Germany). The excitation and detection of eosin was 561 and 580–640 nm, respectively. A 10×/NA 0.4 objective was used with an additional zoom of 1.89. This resulted in a scanned area of 794 × 794 μm in a 1024×1024 pixel format, which led to a pixel size of 0.78 μm. A pinhole setting of 1 Airy was used, resulting in an optical section with an approximate thickness of 5 μm. All images acquired were adapted to the full dynamic range of the system (8 bits). The images were scanned with the epidermis in the horizontal plane of the microscope axis. From each specimen, one section was imaged. For all imaging, the following protocol was used: the image was taken in the middle of the section, directly underneath the epidermis. To minimize the subjective bias associated with the selection of a region of interest, no tissue boundaries were present in the images although the occasional presence of confounding factors such as large blood vessels required user selection.

Distance mapping

Distance Mapping is a newly developed application for morphometry of collagen structures. This measurement method is easily reproducible and can be implemented in existing analysis suites. The average thickness of the collagen bundles and the average distance between the collagen bundles were measured using Qwin Pro. First, segmentation in the grey image was performed, which took place by a delineate correction, in order to isolate the relevant structures (either the bundles or the spaces between the bundles). Then, a binary image was created, containing only the structures of interest (Fig. 1B). The area of interest was selected by both observers, independently from each other. This selection step, however, is not strictly necessary. From the binary image, a distance map (Fig. 1C) as well as a skeleton (Fig. 1D) was created. The distance map is a grey image created from a binary image where for each point the grey value is set equal to its distance to the nearest edge. A skeleton reduces a structure to a pixel-wide line by successive erosion steps lying in the centre of the original structure. The skeleton was used as a mask to measure the distance map (Fig. 1E). The grey value was measured by calculating the distance from the centre of the structure to its edge. Doubling this grey value generates the thickness of the structures representing the *bundle thickness*, in this paper referred to as *bundle thickness*$_{DM}$ and the *distance between (the edges of) the bundles*, also referred to as *bundle spacing*$_{DM}$.

Fourier analysis – first-order maximum

Images were analyzed with the Fast Fourier Transform module of the Qwin Pro image analysis software (version 2.8, Leica Imaging Systems, Cambridge, U.K.). From each image, a first-order maximum power spectrum was generated. Manual thresholding was performed to visualize two clear centres of gravity in the power spectrum. Figure 2 displays the power spectrum of the first-order maximum, where 'A' (Fig. 2D) corresponds to the distance in pixels between the centre of the first-order power spectrum and the centres of gravity of this power spectrum. This distance is used as a measure of the averaged distance between the centres of the collagen bundles, which was called 'bundle packing' in a previously published paper by de Vries *et al.* (2000), In this study, we refer to the averaged distance between the centres of the collagen bundles as the BCD_{FFT}. To determine this distance in micrometres, the following formula was used: [794/A], which is specifically applicable for a scanned area of 794 × 794 μm.

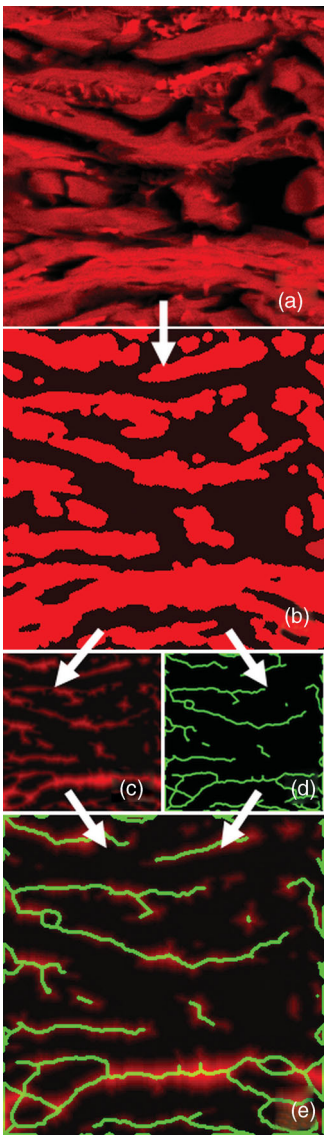

Fig. 1. An explanation of the Distance Mapping method. (A) A fluorescent image representing collagen bundles in healthy skin. These images represent the steps necessary to calculate the bundle thickness. (B) Binary image after segmentation of 'A'. From (B), a distance map (C) as well as a skeleton is generated (D). The skeleton (D) is used as mask to measure the grey intensities in 'C', which are visualized together in 'E'. The measured grey values represent half the distance of the structures in 'B'. Therefore, doubling these grey values is necessary to generate the thickness of the structures.

Statistical analysis

The data were analyzed using SPSS 17.0 for Windows (SPSS Inc., Chicago, IL, U.S.A.). Measurement of the interobserver reliability was based on the measurements of two observers. The *Intraclass Correlation Coefficient (ICC)*, with its 95% confidence interval, was calculated to assess the interobserver reliability for two observers and the reliability of one observer, also known as the average measure ICC and the single measure ICC, respectively (Shrout & Fleiss, 1979). The two-way random effect model was selected and calculated for absolute agreement of the scores. An ICC value of 0.7 was a minimum requirement for reliable results (Nunnally, 1978). We used the single measure ICC for interpretation of our results. The standard error of measurement ($SE_{meas} = \sqrt{}$mean square residual) was used for calculating the amount of error between measurements. Subsequently, the *coefficient of variation (CV)* could be calculated using the following formula: $CV(\%) = [(SE_{meas}/\text{mean}) \times 100]$, whereby a low coefficient of variation represents a better agreement between the observers and a lower measurement error, in comparison with data possessing a high coefficient of variation. It is important to calculate both the reliability (ICC) and the agreement (CV) of a measurement method (de Vet *et al.*, 2006). For normally distributed independent data, statistical testing was performed using the independent *t*-test. This test was used to determine whether both measurement methods were able to measure a clinically important change (Streiner & Norman, 2008a,b). In addition, the responsiveness was calculated by the *effect size* calculated by [(mean $value_{normal\ skin}$ − mean $value_{scar\ tissue}$) / $SD_{normal\ skin}$]. An effect size of 0.2 was considered as small, 0.5 as medium and 0.8 or higher as large (Cohen, 1977; Kazis *et al.*, 1989). A large effect size reflects an excellent responsiveness. Responsiveness can be considered as a part of the validation process of a measurement method and provides information on whether measurement a method is sensitive enough to measure a treatment effect in clinical practice. The two-tailed significance criterion was set at 0.05.

Results

The outcome parameters of Distance Mapping and the Fourier first-order maximum are visualized in Figure 3. This figure shows that the bundle $thickness_{DM}$ and the bundle $spacing_{DM}$ can be considered separately in a straightforward manner, which is more practical and easier to interpret compared to the measure BCD_{FFT}. For example, when a large BCD_{FFT} is calculated one cannot distinguish whether this is due to a large bundle thickness or a large bundle spacing (or a combination of both), whereas in Distance Mapping the bundle thickness and bundle spacing can be easily resolved separately.

The results on the interobserver reliability of Distance Mapping and the Fourier first-order maximum are presented in Table 1. The interobserver reliability was calculated by the single measure ICC. For measurements of the bundle $thickness_{DM}$, the bundle $spacing_{DM}$ and the BCD_{FFT} in healthy skin and scar tissue, the single measure ICC was consistently higher than 0.7, which is a minimum requirement for reliable results: the single measure interobserver reliability for Distance Mapping was above 0.91 for healthy skin and above 0.89 for scar tissue. The single measure interobserver reliability for the Fourier first-order maximum was above 0.88

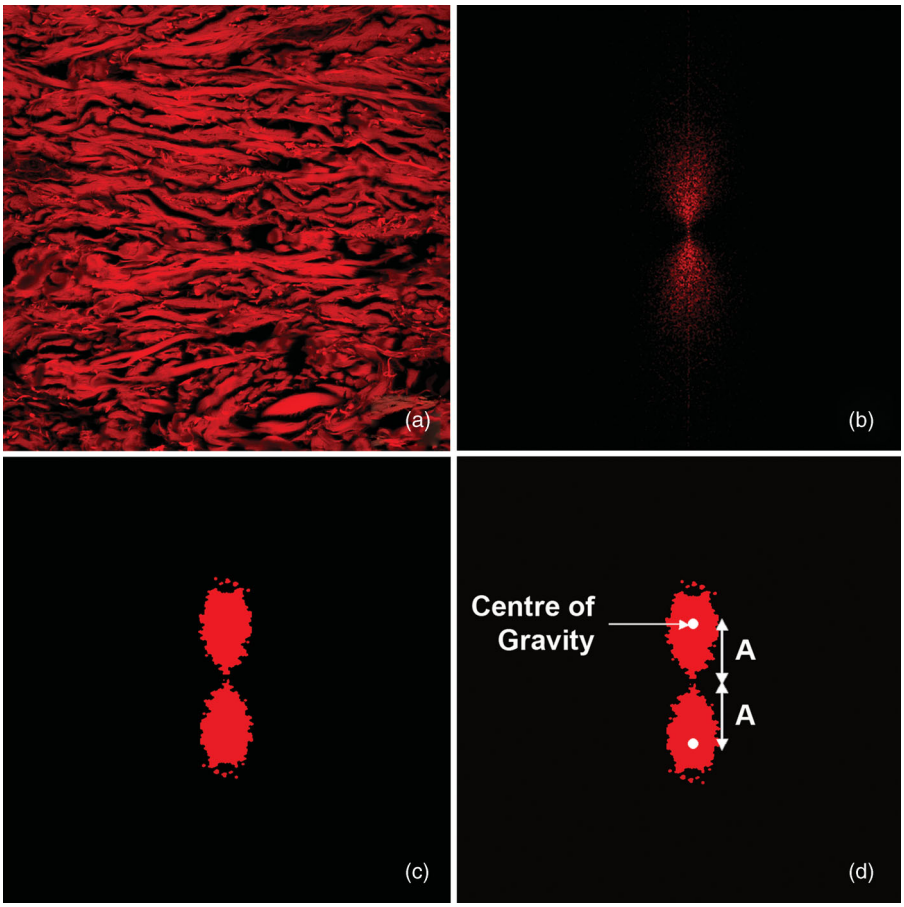

Fig. 2. An explanation of the Fourier first-order maximum method. (A) Fluorescent image of size 794 × 794 μm, representing collagen bundles in healthy skin. (B) Power spectrum of the image presenting the two maxima of the Fourier first-order maximum after manual thresholding. (C) Result after conversion of the power spectrum into a binary image. (D) Explanation of the calculation of the real distance between the centres of the collagen bundles. 'A' = the distance between the centre of the power spectrum and the centre of gravity: BCD_{FFT} (μm) = [794/A].

for healthy skin and above 0.72 for scar tissue. In this study, only the interobserver reliability and not the intraobserver reliability (two repeated measurements of the same observer) was calculated. Because the interobserver reliability was good, separate calculation of the intraobserver reliability, which is less susceptible to variation and bias (because the measurement is performed by the same observer instead of two different observers) was therefore not necessary: the ICCs for the intraobserver reliability would have been even higher.

Results on the coefficients of variation, which are a measure for the agreement between the observers and the measurement error, are displayed in Table 1. It was shown that for measurements on healthy skin, lower coefficients of variation were calculated in Distance Mapping (5.5% and 6.0% for the parameters bundle $thickness_{DM}$ and bundle $spacing_{DM}$, respectively) compared to a higher coefficient of variation for the Fourier first-order maximum BCD_{FFT}, 10.7%. Similarly, low coefficients of variation emerge from scar tissue measurements in Distance Mapping: 2.1% and 6.0% for bundle $thickness_{DM}$ and bundle $spacing_{DM}$, respectively. By contrast, a coefficient of variation of 23.7% was measured for BCD_{FFT} using the Fourier first-order maximum.

Table 2 displays an overview of the mean values of the different outcome parameters for healthy skin and scars. For bundle $thickness_{DM}$, bundle $spacing_{DM}$ and BCD_{FFT} significant differences were found between healthy skin and scar tissue for measurements performed by observer 1, observer 2 and the average of both observers. Relative to scar tissue, healthy skin exhibited significantly thicker collagen bundles, with a significantly larger bundle spacing and a larger BCD_{FFT}. As the bundle $thickness_{DM}$, the bundle $spacing_{DM}$ and the BCD_{FFT} values in healthy skin differed significantly from these values in scar tissue, it can be postulated that both measurement methods are able to measure a clinically important change, viz. the distinction between healthy skin and scar tissue.

To determine the responsiveness for both measurement methods, as a surrogate for validity, the effect size was calculated (Table 2), which provides an indication of whether a measurement method is sensitive enough to measure treatment effect in clinical practice. For Distance Mapping,

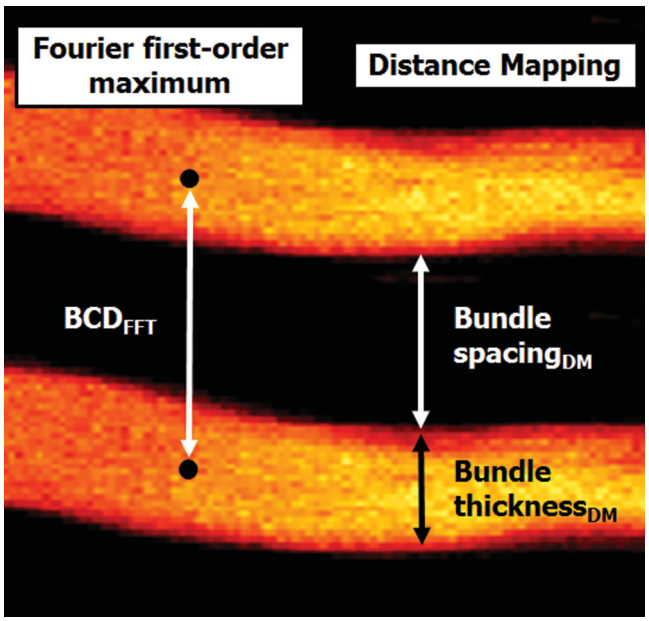

Fig. 3. The visualization of outcome parameters for collagen morphometry as measured by Fourier first-order analysis and Distance Mapping. A reconstructed and enlarged confocal microscopy image of collagen bundles. The fluorescent structures correspond to collagen bundles, which are interspersed with black background representing the space between the collagen bundles. This image displays the outcome parameters for the Fourier first-order maximum (BCD_{FFT}) on the left and for Distance Mapping (bundle $thickness_{DM}$ and bundle $spacing_{DM}$) on the right.

the effect size ranged between 0.481 and 0.619, which can be categorized as a medium effect size. For the Fourier first-order maximum analysis, the effect size ranged between 0.680 and 0.843, which represents a medium to large effect size. Thus, according to standard criteria, these calculated effect sizes reflect a medium to excellent responsiveness (Cohen, 1977; Kazis *et al.*, 1989).

Discussion

In this study, we have critically evaluated two objective measurement methods for collagen bundle morphometry in healthy skin and scars. We introduced a new measurement method: Distance Mapping, which was applied for the first time for collagen morphometry and demonstrated to be a highly reliable method for measuring bundle thickness and bundle spacing in healthy skin and scar tissue. In particular, an almost perfect reliability was observed, which means that only one measurement of one observer is required to obtain highly reliable results. With the Fourier first-order maximum analysis, a good reliability for measuring the average distance between the centres of the collagen bundles (i.e. BCD_{FFT}) in healthy skin and scars was demonstrated: again, only one measurement of one observer is necessary to obtain reliable results.

Besides providing reliable measurements, Distance Mapping and the Fourier first-order maximum were also found to be valid measurement methods, as supported by two important findings in this paper: first, their ability to measure a clinically important change and secondly a good responsiveness. In particular, both methods can detect a clinically important change by distinguishing the collagen morphology in healthy skin from collagen morphology in scar tissue: significantly thicker collagen bundles and significantly more spacing between the collagen bundles were found in healthy skin compared to scar tissue. This discriminative capacity reflects a clinically important change. Furthermore, both methods also exhibited a good responsiveness, that is, a sufficient sensitivity to highlight a treatment effect in clinical practice. In the future, further criterion validation of these methods is recommended, which means that their outcome measures are compared to a 'gold standard' method; however, at present there is no such gold standard.

Preferably, measurement methods should be automated, without necessitating observer interventions. Fourier analysis requires an observer for determining a threshold step in

Table 1. Interobserver reliability of Distance Mapping and the Fourier first-order.

	Distance Mapping		Fourier first-order maximum
	Bundle $thickness_{DM}$	Bundle $spacing_{DM}$	BCD_{FFT}
Healthy skin			
Single measure ICC (95% CI)	0.90 (0.83–0.94)	0.91 (0.84–0.95)	0.88 (0.80–0.93)
Average measure ICC (95% CI)	0.95 (0.91–0.97)	0.95 (0.92–0.97)	0.94 (0.89–0.97)
%CV (SE_{meas})	5.5 (0.26)	6.0 (0.49)	10.7 (1.53)
Scar tissue			
Single measure ICC (95% CI)	0.99 (0.98–0.99)	0.89 (0.80–0.94)	0.72 (0.49–0.85)
Average measure ICC (95% CI)	0.99 (0.99–1.00)	0.94 (0.89–0.97)	0.84 (0.65–0.92)
%CV (SE_{meas})	2.1 (0.09)	6.0 (0.43)	23.7 (2.54)

Note: ICC = Intraclass Correlation Coefficient. For all ICCs, the *p*-value was lower than 0.001. CV = coefficient of variation, SE_{meas} = standard error of measurement and CI = confidence interval.

Table 2. General characteristics of the measured sections of healthy skin and scar tissue.

		Distance Mapping						Fourier first-order maximum		
		Bundle thickness$_{DM}$			Bundle spacing$_{DM}$			BCD$_{FFT}$		
	N	Obs. 1	Obs. 2	Average	Obs. 1	Obs. 2	Average	Obs. 1	Obs. 2	Average
Healthy skin, mean (SD) in μm	50	4.63 (0.81)	4.68 (0.81)	4.65 (0.79)	8.26 (1.70)	8.07 (1.64)	8.17 (1.63)	14.54 (4.27)	13.91 (4.92)	14.22 (4.48)
Scar tissue, mean (SD) in μm	50	4.24 (0.86)	4.28 (0.87)	4.26 (0.86)	7.30 (1.47)	7.06 (1.34)	7.18 (1.37)	11.64 (6.15)	9.76 (4.16)	10.70 (4.93)
95% CI of the difference		0.06–0.72	0.07–0.74	0.07–0.73	0.33–1.59	0.42–1.61	0.39–1.58	0.80–5.00	2.34–5.96	1.66–5.40
p-value		0.022	0.019	0.019	0.003	0.001	0.001	0.007	<0.001	<0.001
Effect size		0.481	0.494	0.500	0.565	0.619	0.603	0.680	0.843	0.787

Note: Mean values of the bundle thickness$_{DM,}$ bundle spacing$_{DM}$ and BCD$_{FFT}$ of the collagen structures in the sections, calculated in micrometres. Statistical testing was performed using the independent *t*-test. CI = confidence interval. Effect size (ES) = [(mean value$_{normal\ skin}$ − mean value$_{scar\ tissue}$) / SD$_{normal\ skin}$]. An ES of 0.2 was considered as small, 0.5 as medium and 0.8 or greater as large. The larger the effect size, the more sensitive the measurement method can be considered to measure a treatment effect in clinical practice.

the image analysis; nonetheless, the interobserver reliability remained good. Distance Mapping, by contrast, does not strictly necessitate an observer intervention. This difference between these measurement methods concerning the observer dependency may be reflected in a difference in the ICCs and the coefficients of variation. First, a better reliability was seen for Distance Mapping compared to Fourier analysis: the ICCs for both healthy and scarred skin, as calculated by Distance Mapping, were higher than the ICCs associated with the Fourier analysis. Second, the relatively low coefficients of variation of Distance Mapping reflected a better agreement between the two observers and a lower measurement error, by contrast to the Fourier analysis, as may be inferred from Table 1.

In conclusion, in this paper, we introduced Distance Mapping and the Fourier first-order maximum as reliable and valid measurement methods to objectively assess collagen bundle morphology. Because of a superior reliability and more useful outcome parameters, Distance Mapping emerges as the preferred and more practical method. Results of this study could contribute to the implementation of these promising measurement techniques into the objective histopathological analyses of healthy skin and scar tissue. Moreover, both these techniques could also be applicable for morphometry of fibrosis of other viscera, such as the heart or the lung.

Acknowledgements

This research was supported by a grant from the Dutch Burns Foundation (number 08.107). We would like to acknowledge Wim Tuinebreijer, MD, PhD for his advice concerning statistical analyses and the researchers of the laboratory of the Association of Dutch Burn Centres for processing the histological slides.

Conflict of interests

The authors state no conflict of interest.

References

Baak, J.P. (2002) The framework of pathology: good laboratory practice by quantitative and molecular methods. *J. Pathol.* **198**, 277–283.

Bloemen, M.C., van der Veer, W.M., Ulrich, M.M., van Zuijlen, P.P., Niessen, F.B. & Middelkoop, E. (2009) Prevention and curative management of hypertrophic scar formation. *Burns* **35**, 463–475.

Chaudhuri, S., Nguyen, H., Rangayyan, R.M., Walsh S. & Frank, C.B. (1987) A Fourier domain directional filtering method for analysis of collagen alignment in ligaments. *IEEE Trans. Biomed. Eng.* **34**, 509–518.

Cohen, J. (1977) *Statistical Power Analysis for the Behavioural Sciences.* Academic Press, New York.

Ferdman, A.G. & Yannas, I.V. (1993) Scattering of light from histologic sections: a new method for the analysis of connective tissue. *J. Invest. Dermatol.* **100**, 710–716.

Guerin, H.A. & Elliott, D.M. (2006) Degeneration affects the fiber reorientation of human annulus fibrosus under tensile load. *J. Biomech.* **39**, 1410–1418.

Har-Shai, Y., Amar, M. & Sabo, E. (2003) Intralesional cryotherapy for enhancing the involution of hypertrophic scars and keloids. *Plast. Reconstr. Surg.* **111**, 1841–1852.

Har-Shai, Y., Sabo, E., Rohde, E., Hyams, M., Assaf, C. & Zouboulis, C.C. (2006) Intralesional cryosurgery enhances the involution of recalcitrant auricular keloids: a new clinical approach supported by experimental studies. *Wound Repair Regen.* **14**, 18–27.

Kazis, L.E., Anderson, J.J. & Meenan, R.F. (1989) Effect sizes for interpreting changes in health status. *Med. Care* **27**, S178–189.

Melis, P., Noorlander, M.L., van der Horst, C.M.A.M. & van Noorden, C.J.F. (2002) Rapid alignment of collagen fibers in the dermis of undermined and not undermined skin stretched with a skin-stretching device. *Plast. Reconstr. Surg.* **109**, 674–681.

Ng, C.P., Hinz, B. & Swartz, M.A. (2005) Interstitial fluid flow induces myofibroblast differentiation and collagen alignment in vitro. *J. Cell. Sci.* **118**, 4731–4739.

Niessen, F.B., Spauwen, P.H., Schalkwijk, J. & Kon, M. (1999) On the nature of hypertrophic scars and keloids: a review. *Plast. Reconstr. Surg.* **104**, 1435–1458.

Noorlander, M.L., Melis, P., Jonker, A. & van Noorden, C.J.F. (2002) A quantitative method to determine the orientation of collagen fibers in the dermis. *J. Histochem. Cytochem.* **50**, 1469–1474.

Nunnally, J.C. (1978) *Psychometric Theory*. McGraw-Hill, New York.

Palmer, B.M. & Bizios, R. (1997) Quantitative characterization of vascular endothelial cell morphology and orientation using Fourier transform analysis. *J. Biomech. Eng.* **119**, 159–165.

Shrout, P.E. & Fleiss, J.L. (1979) Intraclass correlations: uses in assessing rater reliability. *Psychol. Bull.* **86**, 420–428.

Singer, A.J. & McClain, S.A. (2006) Development of a porcine incisional wound model and novel scarring scales. *Wound Repair Regen.* **14**, 492–497.

Streiner, D.L. & Norman, G.R. (2008a) Reliability. *Health Measurement Scales: a practical guide to their development and use* (ed. by D.L. Streiner, G.R. Norman), pp. 167–207. Oxford University Press, Oxford.

Streiner, D.L. & Norman, G.R. (2008b) Validity. *Health Measurement Scales: a practical guide to their development and use* (ed. by D.L. Streiner, G.R. Norman), pp. 247–276. Oxford University Press, Oxford.

Verhaegen, P.D., van Zuijlen, P.P., Pennings, N.M., van Marle, J., Niessen, F.B., van der Horst, C.M. & Middelkoop, E. (2009) Differences in collagen architecture between keloid, hypertrophic scar, normotrophic scar, and normal skin: an objective histopathological analysis. *Wound Repair Regen.* **17**, 649–656.

de Vet, H.C., Terwee, C.B., Knol, D.L. & Bouter, L.M. (2006) When to use agreement versus reliability measures. *J. Clin. Epidemiol.* **59**, 1033–1039.

de Vries, H.J., Enomoto, D.N., van Marle, J., van Zuijlen, P.P., Mekkes, J.R. & Bos, J.D. (2000) Dermal organization in scleroderma: the fast Fourier transform and the laser scatter method objectify fibrosis in nonlesional as well as lesional skin. *Lab Invest.* **80**, 1281–1289.

Whittaker, P., Boughner, D.R. & Kloner, R.A. (1991) Role of collagen in acute myocardial infarct expansion. *Circulation* **84**, 2123–2134.

Whittaker, P., Kloner, R.A., Boughner D.R. & Pickering, J.G. (1994) Quantitative assessment of myocardial collagen with picrosirius red staining and circularly polarized light. *Basic Res. Cardiol.* **89**, 397–410.

van Zuijlen, P.P., Lamme, E.N., van Galen, M.J., van Marle, J., Kreis, R.W. & Middelkoop, E. (2002a) Long-term results of a clinical trial on dermal substitution. A light microscopy and Fourier analysis based evaluation. *Burns* **28**, 151–160.

van Zuijlen, P.P., de Vries, H.J., Lamme, E.N., Coppens, J.E., van Marle, J., Kreis, R.W. & Middelkoop, E. (2002b) Morphometry of dermal collagen orientation by Fourier analysis is superior to multi-observer assessment. *J. Pathol.* **198**, 284–291.

van Zuijlen, P.P. *et al.* (2003) Collagen morphology in human skin and scar tissue: no adaptations in response to mechanical loading at joints. *Burns* **29**, 423–431.