
Reliability of widefield capillary microscopy to measure nailfold capillary density in systemic sclerosis

M. Hudson, A. Masetto, R. Steele, E. Arthurs, M. Baron
for the Canadian Scleroderma Research Group*

Division of Rheumatology, Jewish General Hospital and McGill University, Montreal, Quebec, Canada; Division of Rheumatology, University of Sherbrooke, Sherbrooke, Quebec, Canada.

Marie Hudson, MD, MPH
Ariel Masetto, MD
Russell Steele, PhD
Erin Arthurs, BSc
Murray Baron, MD

*Investigators of the Canadian Scleroderma Research Group: see Appendix for list.

This study was funded by an Établissement de Jeunes Chercheurs award from the Fonds de la Recherche en Santé du Québec. Dr Hudson is supported by a New Investigator Award from the Canadian Institutes of Health Research.

Please address correspondence and reprint requests to:

Dr Marie Hudson,
Jewish General Hospital, Room A-216,
3755 Cote Ste Catherine Street,
Montreal, Quebec H3T 1E2, Canada.
E-mail: marie.hudson@mcgill.ca

Received on February 4, 2010; accepted in revised form on September 27, 2010.

Clin Exp Rheumatol 2010; 28 (Suppl. 62): S36-S41.

© Copyright CLINICAL AND EXPERIMENTAL RHEUMATOLOGY 2010.

Key words: systemic sclerosis, widefield microscopy, reliability

ABSTRACT

Objectives. To determine intra- and inter-observer reliability of widefield microscopy to measure nailfold capillary density in patients with systemic sclerosis (SSc).

Methods. Five SSc patients were examined with a STEMV-8 Zeiss biomicroscope with 50x magnification. The nailfold of the second, third, fourth and fifth fingers of both hands of each patient were photographed twice by each of two observers, once in the morning and again in the afternoon (total of 32 pictures). Two raters reviewed the photographs to produce capillary density readings. Intra- and inter-rater reliability of the readings were computed using intra-class correlations (ICC). Additional analyses were undertaken to determine the impact of other sources of variability in the data, namely patient, finger, technician and time.

Results. Intra- and inter-rater reliability were substantial (ICC 0.72-0.84) when raters were reading the same photographs or photographs taken at the same time of day. Agreement was only fair between morning and afternoon density readings (ICC 0.30-0.37). Patients, individual fingers and technician accounted for a large part of the variability in the data (combined variance component of 7.69 out of the total 12.23). The coefficient of variation of widefield microscopy was 24%.

Conclusion. Although intra- and inter-rater reliability of nailfold capillary density measurements using widefield microscopy are good, proper standardisation of the conditions under which capillaroscopy is done and better imaging of nailfold capillary abnormalities should be considered if nailfold capillary density is to be used as an outcome measure in multi-centre clinical trials in SSc.

Introduction

Systemic sclerosis (SSc) is a chronic inflammatory disease characterised by thickening and fibrosis of the skin and internal organs, most commonly the gastrointestinal tract, lungs, and heart. Vascular disease is ubiquitous in SSc and is manifested by Raynaud's phenomenon, digital ulcers, telangiectasias, scleroderma renal crisis and pulmonary hypertension.

The current classification criteria for SSc are preliminary criteria developed by the American College of Rheumatology in 1980 (1). They include skin thickening, sclerodactyly, digital pits or loss of substance of the distal finger pads, and/or pulmonary fibrosis. However, it is well-known that these criteria lack sensitivity especially for patients with limited disease (2). Updated criteria including nailfold capillary abnormalities have been proposed (3) and shown to improve the sensitivity of the classification criteria (4, 5).

Several methods exist to examine nailfold capillaries in SSc, of which widefield microscopy is one of the most well-established techniques (6). The main abnormalities described in SSc include dilated and giant capillary loops, haemorrhages and loss of capillaries and/or avascular areas (7). Abnormalities on widefield microscopy have been shown to be of considerable clinical importance, among other things for the diagnosis of SSc (7) and prognosis of patients with Raynaud's phenomenon (8). To date, most of the studies using widefield microscopy have been qualitative, although some have been semi-quantitative or included measurement of capillary dimensions and density (9-11). However, there is a paucity of information on the reliability of these assessments.

Novel therapies, involving angiogenesis and vasculogenesis, are currently

Competing interests: none declared.

being investigated to treat the vasculopathy in SSc (12, 13). In those settings, density, besides morphology, of nailfold capillaries may become a useful outcome measure. There are no studies that have assessed the reliability of density measurements of nailfold capillary abnormalities in SSc using widefield microscopy.

Given that reproducible measures of nailfold capillaries are needed for future therapeutic trials in SSc, we undertook this study to determine the intra- and inter-observer reliability of widefield microscopy to measure nailfold capillary density in SSc patients.

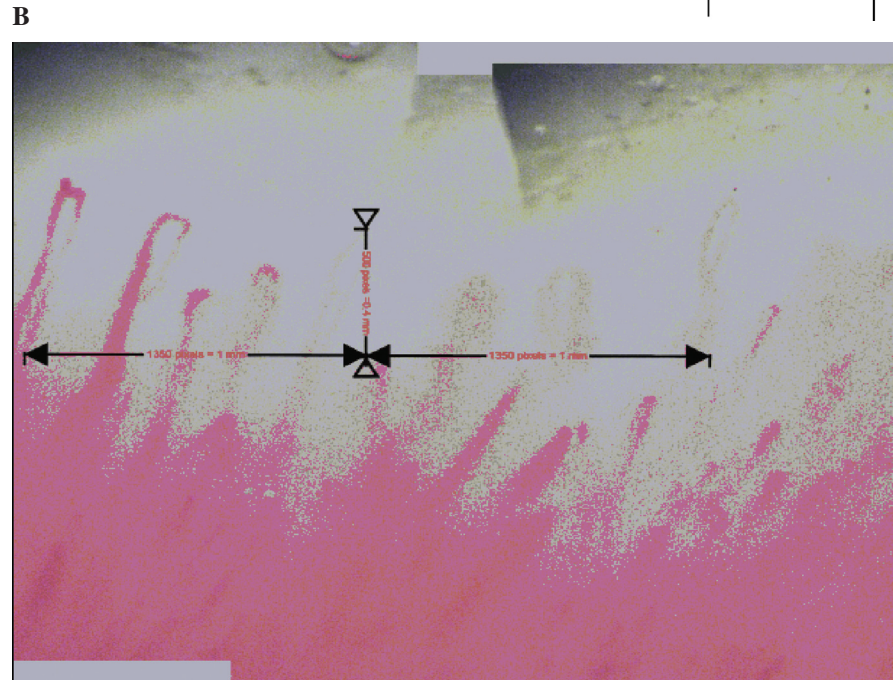
Methods

Five SSc patients from the Canadian Scleroderma Research Group registry with records indicating that they had nailfold capillary abnormalities were identified and invited to participate in this study. Ethics approval for the study and informed written patient consents were obtained.

Two investigators (MH, AM) performed the examinations, took the photographs and rated the mosaics. We used a STEMV-8 Zeiss biomicroscope with 50x magnification, calibrated to allow transforming pictures measured in pixels to micra and a microscope-integrated digital camera with 10 megapixels resolution. Immersion oil was applied to the nailfold to enhance the transparency of the skin.

Subjects were asked to arrive at least 15 minutes before their exam was scheduled and to wait in a room with a temperature of 20–22°C. The nailfold (distal row) of the second, third, fourth and fifth fingers of both hands of each patient (total of 8 fingers) were examined and photographed by two observers, once by each observer in the morning and again by each observer in the afternoon (total of 4 examinations). Patients were positioned for the morning and afternoon examinations and seen by each observer consecutively. Thus, there was no re-positioning between the two morning examinations and the two afternoon examinations. At each examination and for each finger, each observer took photos of three fields (approximately one linear millime-

Fig. 1. A. Cartoon of fingertip showing sites where pictures were taken, in the centre of the nailfold and left and right; **B.** Example of a mosaic used to read capillary density (showing capillary density of 8 loops/1700 pixels).



tre each) in the middle of the nailfold and on either side (Fig. 1A). Pictures were assembled into mosaics using the software package from Photoshop CS 2 (Fig. 1B). The order of patients and the 32 mosaics of each patient were randomly arranged (using computer generated random numbers) prior to sending the mosaics to the 2 raters for capillary density measurements.

The 2 raters reviewed all the mosaics to produce capillary density readings. Capillary density was defined as the number of capillaries in the distal capillary row of the nailfold of each digit over a pre-defined distance. Specifically, a sentinel loop was identified (Fig. 1B). A vertical line measuring 500 pixels from the top of the sentinel loop was drawn. Then, a horizontal line measuring 1350 pixels on both sides of the vertical line was drawn. The total length used for measurements was thus 2700 pixels, equivalent to 2000 micra or 2mm. All complete and partial capillary loops that crossed the horizontal line were included in the count.

Intra- and inter-rater reliability of the readings were computed using intra-class correlations (ICC) and weighted Kappas. Agreement was interpreted using the following guidelines: 0=poor, 0.01–0.20=slight, 0.21–0.40=fair, 0.41–0.60=moderate, 0.61–0.80=substantial, and 0.81–1=almost perfect (14).

In addition to inter-rater variability, we identified 4 other possible sources of variability in measuring capillary density: the technician responsible for acquiring the photos (in this case, also raters 1 and 2), the timing of the exam (in this case the morning and afternoon), the finger being photographed (in this study fingers 2–5 and excluding the thumb) and the patient. We therefore undertook additional analyses to determine the impact of these other sources of variability on the measurements. Exploratory analyses using fixed effect Analysis of Variance were conducted to examine the relationship between variables and address possible interactions between patients, fingers and time. More formal analyses using linear

mixed models including both random effects terms and fixed effects terms were then used to evaluate the variance components of each variable and to account for nesting of the potential sources of variability. Mixed models are particularly useful for the kind of clustered or dependent data that we collected in our study. ICCs were re-calculated using the variance components from the mixed models to estimate the reliability of the raters based on the new classification of the subjects.

Like all quantitative tests in clinical medicine, widefield microscopy is not perfectly reproducible even when performed in exact accordance with standard procedures every time. Precision is an attribute of a quantitative measurement technique that refers to the ability to reproduce the same numerical result in the setting of no real biologic change when the test is repeatedly performed in an identical fashion. We evaluate precision by using the percent coefficient of variation (%CV), calculated by dividing the standard deviation by the mean (and multiplying by 100). Using the results of the regression analysis, we also computed the raw precision error of widefield microscopy.

The funding sources had no role in the design of the study, analysis of the data, preparation of the manuscript and decision to submit for publication.

Results

Characteristics of the cohort

This study included 5 patients. Mean age (standard deviation) of the patients was 45 (± 18) years; there were 2 women and 3 men; mean duration of disease measured from the time of onset of the first non-Raynaud's manifestation of SSc was 4.3 (± 2) years; all had limited scleroderma. Mean nailfold capillary density was 8.86 (± 4.15), equivalent to 4.43 loops/mm.

Inter-rater reliability

Intra-rater reliability was assessed by having rater 1 measure the capillary density of the 32 mosaics from one study patient twice, 5 months apart. The ICC on the two sets of 32 readings was 0.79, indicating substantial intra-rater agreement.

Fig. 2. Scenarios in which inter-rater reliability was assessed

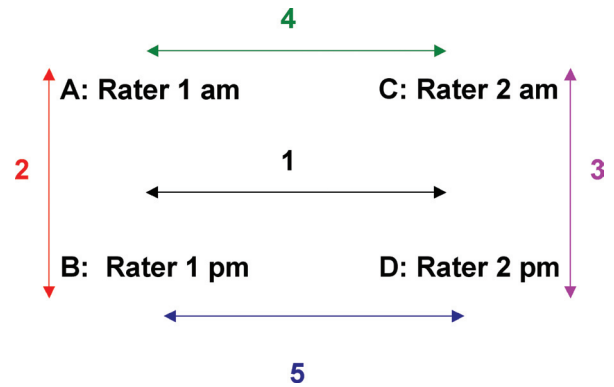


Table I. Inter-rater reliability statistics between capillary density readings of two raters, from two sets of photos taken in the morning and afternoon.

Scenarios	1 R1 vs R2 (all photos, am & pm)	2 R1 am vs. R1 pm (R1 photos)	3 R2 am vs. R2 pm (R2 photos)	4 R1 am vs. R2 am (am photos)	5 R1 pm vs. R2 pm (pm photos)
ICC	0.74	0.30	0.34	0.74	0.73
Weighted Kappa	0.80	0.30	0.37	0.84	0.72

R1: Density readings by rater 1 (MH); R2: Density readings by rater 2 (AM); ICC: intra-class correlation.

Figure 2 and Table I display the inter-rater agreement statistics for the following cases: 1) agreement between rater 1 density readings and rater 2 density readings for all photos (for every finger at every time); 2) agreement between rater 1 density readings of rater 1 photos in the morning and rater 1 photos in the afternoon; 3) agreement between rater 2 density readings of rater 2 photos in the morning and rater 2 photos in the afternoon; 4) agreement between rater 1 density readings of all morning photos and rater 2 density readings of all morning photos; and, 5) agreement between rater 1 density readings of all afternoon photos and rater 2 density readings of all afternoon photos. The results show at least substantial agreement between raters reading the same photographs or between photographs taken at the same time of day (ICC 0.72–0.84). There was only fair agreement between morning and afternoon density readings (ICC 0.30–0.37). As the inter-rater agreement on the actual photographs was high, this finding suggests a true difference in nailfold density depending on fluctuating physiological and environmental factors.

The intra-class correlations between

rater 1 and rater 2 density readings for all photos *per finger* were also calculated. The statistics show that agreement between raters, all be it substantial, was the lowest for finger 2 (index finger, ICC 0.66) and highest for finger 5 (little finger, ICC 0.82). There was also substantial agreement for fingers 3 and 4 (ICC 0.77 and 0.73, respectively).

Sources of variability

In addition to inter-rater variability, we identified 4 other possible sources of variability in measuring capillary density: the technician responsible for acquiring the photos (in this case, also raters 1 and 2), the timing of the exam (in this case the morning and afternoon), the finger being photographed (in this study fingers 2–5 and excluding the thumb) and the patient. An exploratory analysis revealed a significant interaction between patient and fingers, suggesting structuring the data with fingers as the subjects of analysis as opposed to the patients (Table II). It also revealed a significant interaction between time of study and patient, consistent with the finding that inter-rater reliability was only fair at different time points.

Table II. Analysis of Variance model used to explore how to structure the data for the regression analysis. The results suggest that patient finger rather than patient should be the units of analysis.

Source	Degrees of Freedom	Sums of Squares	Mean Square	F value	p-value
Finger	7	577.76	82.54	16.57	< 2.2e-16
Time	1	38.37	38.37	7.70	0.006
Patient	3	292.92	97.64	19.60	4.313e-11
Finger: Time	7	25.91	3.70	0.74	0.636
Finger: Patient	21	763.24	36.34	7.29	2.277e-15
Time: Patient	3	88.31	29.44	5.90	0.001
Patient: Finger:Time	21	221.39	10.54	2.12	0.004
Residuals	187	931.67	4.98		

Table III. Mixed regression model for identifying sources of variability in the data.

Random effects	Variance component	Standard Deviation
Patient-finger-time	1.088	1.043
Patient-finger	5.228	2.286
Patient	0.726	0.852
Photo	0.518	0.720
Rater	0.132	0.363
Residual	4.539	2.130

Table IV. Intraclass correlations (ICC) generated using the variance components of the mixed effects model.

	ICC
Reliability of density readings between raters on particular subject	0.51
Reliability of density readings between raters on particular subject's photo	0.55
Reliability of density readings between raters on particular subject's photo at particular time of day	0.57
Reliability between photos of particular subject with same rater at same time	0.54

Various mixed models were considered to try to explain the variability in the data, considering the possible sources of variability identified as well as the possible nestedness among the sources. The best fit model was a random effects model, where all variables were considered random effects (Table III). As could be expected, the patient-finger variable accounted for the largest amount of variability, with a variance component of 5.23 out of the total 12.23. The patient-finger-time variable also accounted for an important additional portion of the variability with a variance component of 1.09. On the other hand, the rater, with the lowest variance component of 0.13, contributed very little to the total variability, consistent with our results indicating good inter-rater re-

liability. Although we initially found some variability between photos, we were able to isolate the source for this variability as being patient 2. Review of this patient's mosaics revealed highly disorganised nailfold capillary architecture. Once this outlying patient was removed, photo variability was much lower (variance component 0.52) and the residual variance component decreased twofold. Despite the patient-finger variable accounting for a large amount of the variability in this model, the patient variable alone was included to adjust for possible heterogeneity of the overall mean density for the patients. The variance component for the patient provided some evidence that patients have differences in overall density, but that was dwarfed by the variability within

their own fingers (between subject variance of 0.73 vs. between finger variance of 5.23).

Of note, after accounting for the above sources of variability, the variance component of the residual was still quite large (4.54). One way of interpreting this term is to look at its standard deviation and to compute the coefficient of variation, which represents the precision error in the technique used. The mean overall capillary density was 8.86 and the standard deviation of the residual term was 2.13, resulting in a coefficient of variation of 24%. This suggests considerable precision error in the technique used.

ICCs were re-calculated using the variance components of the mixed effects model to evaluate the overall inter-rater reliability of nailfold capillary microscopy (Table IV). The ICCs were lower than previous estimates that were calculated using only subsets of the data and indicated moderate agreement between raters (range 0.51 to 0.57). This analysis more accurately reflects the agreement between raters for a particular subject's finger, as it properly accounts for all sources of variability that were not accounted for in the analysis of the previous section.

Discussion

This study provides evidence of reasonable intra- and inter-rater reliability for the measurements of nailfold capillary densities of SSc patients using widefield microscopy. In addition, detailed analysis allowed us to determine that most of the structured variability in capillary density measurements comes from patients themselves, with most of the variability being among a patient's own fingers. The acquisition of the photos seemed to account for little of the variability. However, time did, with photos taken at different times of the day (*i.e.* morning and afternoon) having greater variability than photos taken at the same time of day. Finally, over and above all these sources of variability, there was significant independent measurement error, in large part likely resulting from the technique of widefield microscopy used to acquire the images.

Previous studies have measured nailfold capillary density in SSc patients and have reported it to range from 4.7 to 6 loops per millimetre (10, 11, 15). Although lower than those estimates, our finding of 4.4 loops per millimetre is nevertheless consistent with those previous reports. However, although widefield microscopy to measure nailfold capillaries in SSc has been in use for many decades, little data has been published on the intra- and inter-rater reliability of this technique to date. In addition, two studies have considered reliability of capillaroscopy but they used a different method of imaging, namely video-capillaroscopy, and reliability was measured between raters reviewing *one set of photographs* (16, 17). This is the first to study inter-rater reliability on *different photographs* acquired by different technicians using widefield microscopy. We showed that different technicians account for only a small part of the variability in the overall data.

Despite the importance and continued use of widefield microscopy in SSc research (8, 18), the level of magnification is low (in most cases less than 50x) and it requires oil to be applied on the distal nailfold. Hence, poor resolution at that level of magnification and light and oil artefact likely interfere with the precision of the technique. Not surprisingly, we showed that the precision error in the technique itself was considerable. New imaging techniques allowing higher magnification (up to 300x) and better resolution of images have been described (15, 16). Use of this more sophisticated equipment will likely lead to less precision error and better inter-rater reliability in the future.

A limitation of this study is that we did not assess the morphology of nailfold capillaries. Morphological abnormalities such as ramified capillaries have been shown to be useful in the diagnosis and prognosis of SSc, and are currently being investigated in the setting of neoangiogenesis. Nevertheless, the question of interest in this study was to assess the reliability of widefield microscopy as an outcome measure to determine capillary density. Moreover, to the extent that we found that the tech-

nique used lacked precision, further studies are now planned to determine the validity of measuring both density and morphology of nailfold capillaries using more sophisticated equipment with greater magnification and better resolution.

The results of this study are especially important in terms of validating nailfold capillaroscopy as an outcome measure in interventional studies of scleroderma that may require repeated nailfold capillaroscopic examinations. The important points to draw from this analysis are the following: first, inter-rater reliability of nailfold capillaroscopy is good but is reduced by the lack of precision of the imaging technique used; there is little structured variability arising from different technicians acquiring photos of nailfold capillaries, thus supporting multi-centre studies using nailfold capillaroscopy as an endpoint; and, the fact that fingers and time contributed significantly to the variability in the data highlights the importance of standardizing the conditions under which the images are acquired (*e.g.* position, time of day and examination setting for each patient). Failure to use reliable techniques may lead to results that are impossible to interpret (*e.g.* does nailfold capillary density change in response to an intervention or simply because of differences in examination technique). In studies where capillaroscopy may need to be repeated over time, one way of ensuring that the same areas are photographed may be to display previous images either as prints or as computer images to guide the technician to the same region as previously imaged (16). Others have also recently recognised the importance of nailfold capillaroscopy in SSc (19) and the need to standardise the technique (20).

In conclusion, intra- and inter-rater reliability of nailfold capillary density measurements using widefield microscopy is good. However, proper standardization of the conditions under which capillaroscopy is done and better imaging of nailfold capillary abnormalities should be considered if nailfold capillary density is to be used as an outcome measure in multi-centre clinical trials in SSc.

Appendix

Investigators of the Canadian Scleroderma Research Group

M. Baron, Montreal, Quebec
J. Pope, London, Ontario
J. Markland, Saskatoon, Saskatchewan
D. Robinson, Winnipeg, Manitoba
N. Jones, Edmonton, Alberta
N. Khalidi, Hamilton, Ontario
P. Docherty, Moncton, New Brunswick
E. Kaminska, Hamilton, Ontario
A. Masetto, Sherbrooke, Quebec
D. Smith, Ottawa, Ontario
E. Sutton, Halifax, Nova Scotia
J-P. Mathieu, Montreal, Quebec
M. Hudson, Montreal, Quebec
S. Ligier, Montreal, Quebec
T. Grodzicky, Montreal, Quebec
S. Mittoo, Winnipeg, Manitoba
M. Fritzler, Advanced Diagnostics Laboratory, Calgary, Alberta

References

1. SUBCOMMITTEE FOR SCLERODERMA CRITERIA OF THE AMERICAN RHEUMATISM ASSOCIATION DIAGNOSTIC AND THERAPEUTIC CRITERIA COMMITTEE: Preliminary criteria for the classification of systemic sclerosis (scleroderma). Subcommittee for scleroderma criteria of the American Rheumatism Association Diagnostic and Therapeutic Criteria Committee. *Arthritis Rheum* 1980; 23: 581-90.
2. VALENTINI G, BLACK C: Systemic sclerosis. *Best Pract Res Clin Rheumatol* 2002; 16: 807-16.
3. LEROY EC, MEDSGER TA: Criteria for the classification of early Systemic Sclerosis. *Rheumatol* 2001; 28: 1573-6.
4. LONZETTI LS, JOYAL F, RAYNAULD JP *et al.*: Updating the American College of Rheumatology preliminary classification criteria for systemic sclerosis: addition of severe nailfold capillaroscopy abnormalities markedly increases the sensitivity for limited scleroderma. *Arthritis Rheum* 2001; 44: 735-6.
5. HUDSON M, TAILLEFER S, STEELE R *et al.*: Improving the sensitivity of the American College of Rheumatology classification criteria for systemic sclerosis. *Clin Exp Rheumatol* 2007; 25: 754-7.
6. MARICQ HR, LEROY EC: Patterns of finger capillary abnormalities in connective tissue disease by "wide-field" microscopy. *Arthritis Rheum* 1973; 16: 619-28.
7. CUTOLO M, SULLI A, SECCHI ME, PAOLINO S, PIZZORNI C: Nailfold capillaroscopy is useful for the diagnosis and follow-up of autoimmune rheumatic diseases. A future tool for the analysis of microvascular heart involvement? *Rheumatology* (Oxford, England) 2006; 45 (Suppl. 4): iv43-iv6.
8. KOENIG M, JOYAL F, FRITZLER MJ *et al.*: Autoantibodies and microvascular damage are independent predictive factors for the progression of Raynaud's phenomenon to systemic sclerosis: a twenty-year prospec-

- tive study of 586 patients, with validation of proposed criteria for early systemic sclerosis. *Arthritis Rheum* 2008; 58: 3902-12.
9. HERRICK A: Diagnosis and management of scleroderma peripheral vascular disease. *Rheum Dis Clin North Am* 2008; 34: 89-114; vii.
 10. KABASAKAL Y, ELVINS DM, RING EF, MCHUGH NJ: Quantitative nailfold capillaroscopy findings in a population with connective tissue disease and in normal healthy controls. *Ann Rheumatic Dis* 1996; 55: 507-12.
 11. WILDT M, HESSELSTRAND R, AKESSON A, SCHEJA A: Simple counting of nailfold capillary density in suspected systemic sclerosis - 9 years' experience. *Scand J Rheumatol* 2007; 36: 452-7.
 12. ABOU-RAYA A, ABOU-RAYA S, HELMIH M. STATINS: potentially useful in therapy of systemic sclerosis-related Raynaud's phenomenon and digital ulcers. *J Rheumatol* 2008; 35: 1801-8.
 13. DISTLER JH, ALLANORE Y, AVOUAC J *et al.*: EULAR Scleroderma Trials and Research group statement and recommendations on endothelial precursor cells. *Ann Rheum Dis* 2009; 68: 163-8.
 14. LANDIS JR, KOCH GG: The measurement of observer agreement for categorical data. *Biometrics* 1977; 33: 159-74.
 15. CUTOLO M, PIZZORNI C, SECCHI ME, SULLI A: Capillaroscopy. *Best Pract Res Clin Rheumatol* 2008; 22: 1093-108.
 16. ANDERSON ME, ALLEN PD, MOORE T, HILLIER V, TAYLOR CJ, HERRICK AL: Computerized nailfold video capillaroscopy--a new tool for assessment of Raynaud's phenomenon. *J Rheumatol* 2005; 32: 841-8.
 17. SEBASTIANI M, MANFREDI A, COLACI M *et al.*: Capillaroscopic Skin Ulcer Risk Index: A new prognostic tool for digital skin ulcer development in systemic sclerosis patients. *Arthritis Rheum* 2009; 61: 688-94.
 18. MATUCCI-CERINIC M, STEEN V, NASH P, HACHULLA E: The complexity of managing systemic sclerosis: screening and diagnosis. *Rheumatology* (Oxford, England) 2009; 48 (Suppl. 3): iii8-13.
 19. RICCIERI V, GERMANO V, ALESSANDRI C *et al.*: More severe nailfold capillaroscopy findings and anti-endothelial cell antibodies. Are they useful tools for prognostic use in systemic sclerosis? *Clin Exp Rheumatol* 2008; 26: 992-7.
 20. DE ANGELIS R, CUTOLO M, SALAFFI F, RESTREPO JP, GRASSI W: Quantitative and qualitative assessment of one rheumatology trainee's experience with a self-teaching programme in videocapillaroscopy. *Clin Exp Rheumatol* 2009; 27: 651-3.