

EXCELLENT RELIABILITY OF SEMI-QUANTITATIVE NAILFOLD CAPILLAROSCOPY IN PATIENTS WITH SYSTEMIC SCLEROSIS – A PILOT STUDY

Ana Maria Gheorghiu^{1,2}, Raida Oneata¹, Mihai Bojinca^{1,2}, Rucsandra Dobrota^{1,3}, Alina Soare^{1,4}, Liviu Macovei^{1,2}, Mihaela Milicescu^{1,2}, Mariana Sasu², Marilena Gorga¹, Roxana Sfrent-Cornateanu¹, Victor Stoica^{1,2}, Carina Mihai^{1,2}

¹Carol Davila University of Medicine and Pharmacy, Bucharest, Romania

²Cantacuzino Clinical Hospital, Internal Medicine and Rheumatology Department, Bucharest, Romania

³Department of Rheumatology, Zurich University Hospital, Zurich, Switzerland

⁴Department of Internal Medicine III, Institute for Clinical Immunology, University of Erlangen-Nuremberg, Erlangen, Germany

Abstract

Background. Semi-quantitative nailfold capillaroscopy (NFC) scoring represents a promising tool for assessing disease activity, severity and change in systemic sclerosis (SSc), however there is no consensus yet over which capillaroscopy abnormalities should be analyzed and how.

Objective. Investigation of the reliability of the qualitative and semi-quantitative scoring of NFC assessment between two raters and test-retest for each rater in a SSc cohort.

Methods. This is a single-center pilot study where 2 raters assessed the NFC images of 48 consecutive patients with SSc. Data were analyzed in 3 ways: **1.** qualitatively by “normal”/“abnormal” category; **2.** qualitatively by the following categories: “early”, “active”, “late” SSc patterns, “normal”, and unclassifiable in any pattern, and step; **3.** Semi-quantitatively by calculating the mean score for capillary loss, disorganization of the microvascular array, giant capillaries, microhaemorrhages and capillary ramifications and combinations of giant capillaries and microhaemorrhages (as a surrogate for vascular activity). Disorganization and ramifications (surrogate for vascular damage) were also assessed. Variables for all steps were calculated for all fingers and for each finger. Inter-rater/intra-rater agreement was assessed by Cohen's kappa coefficients for qualitative variables and by intraclass correlation coefficients (ICC) for mean score values of abnormalities.

Results. Inter-rater reliability ranged from good to excellent agreement for mean score values of abnormalities in all fingers (ICC coefficients 0.745 to 0.897) and was excellent for activity (ICC coefficient of 0.923) and damage combinations (ICC coefficient of 0.918). Assessment of abnormalities in a qualitative manner (normal/abnormal or with capillaroscopy patterns) showed weaker inter-rater agreement than the semi-quantitative assessment (k coefficient <0.7). Intra-rater variability was good to excellent for mean score values of abnormalities and activity and damage combinations in all fingers and separate fingers for both raters; for qualitative assessment, only one of the raters had good test-retest reliability.

Conclusion. Reliability of NFC assessment is essential in SSc trials/clinical practice to ensure quality of data. This pilot study demonstrates very good reliability between raters of the semi-quantitative NFC assessment in a SSc cohort. Combinations of NFC abnormalities had very good reliability and might be preferred because they are less time consuming.

Keywords: systemic sclerosis, capillaroscopy, agreement, reliability, quantitative

INTRODUCTION AND OBJECTIVE

Systemic sclerosis (SSc) is a systemic autoimmune disease characterized by disseminated, obliterative microvascular changes, followed by fibrosis. Nailfold capillaroscopy (NFC) is a non-invasive method by which SSc-specific microvascular abnormalities can be observed (1). As these changes occur

early in the natural course of the disease, their observation is important for an early diagnosis. SSc-specific NFC-changes are the presence of very enlarged capillaries (giant capillaries) and the progressive decrease of nailfold capillary density; the SSc-specific NFC pattern has been described by Rand-Maricq and co-workers more than four decades ago (2). At

Correspondence address:

Ana Maria Gheorghiu, MD, PhD student, EUSTAR center 100, Cantacuzino Hospital, Carol Davila University of Medicine and Pharmacy, 5-7 Ion Movila Street, Bucharest, Romania.

E-mail: ana.gheghe@gmail.com

the present, these changes are included among the ACR/EULAR classification criteria for SSc (3), whose sensitivity has been shown to be superior to the 1980 ACR classification criteria (4,5). NFC also contributes to identifying, among patients with Raynaud's phenomenon, those who are at risk for developing scleroderma-spectrum diseases (6,7). While in early stages of these diseases microhemorrhages are frequently found on NFC, late disease is characterized by important capillary loss resulting in avascular areas and signs of ineffective vasculogenesis, represented by ramified or bizarre capillaries (8). Several studies have shown that NFC abnormalities correlate with disease activity and severity (9,10). More importantly, successful treatment has been shown to diminish NFC abnormalities in severe SSc cases (11,12).

Semi-quantitative capillaroscopy represents a promising tool for assessing disease activity, severity and change in SSc, however there is no consensus yet over which capillaroscopic abnormalities should be analyzed and how (13). The aim of this single-center study was to investigate the reliability of the qualitative and semi-quantitative scoring of NFC assessment between two raters and test-retest for each rater in a cohort of patients with SSc.

PATIENTS AND METHODS

Forty-eight consecutive patients with SSc attending our EUSTAR center between Aug. 2015 – Jan. 2016, all satisfying the ACR 1980 or the ACR/EULAR 2013 classification criteria (3,4), were included in this study. Sixteen healthy persons, selected among our clinic's staff, were also examined. Patients had at least 15 min. to accommodate to room temperature (22-25°C). Widefield NFC was performed with a SZX 7 Olympus stereo-microscope, using an external cold light source. Fingers 2 to 5 of both hands were examined, after placing a drop of immersion oil of each cuticle. Magnification was X 40 and for each finger, the image with most abnormalities was recorded and stored, using the software QuickPHOTO MICRO 3.0 (Promicra, Prague, Czech Republic).

Two raters independently reviewed the NFC images and assessed the following abnormalities: giant capillaries, microhemorrhages, ramified capillaries, disorganization and capillary density per linear millimeter. Each abnormality was noted as present/ab-

sent (qualitative analysis) and also evaluated semi-quantitatively on a scale from 0 to 3 where: 0=absent, 1=<33% alteration, 2=33-66% alteration and 3=>66% alteration (14). Qualitative NFC was completed by classifying each patient as having "normal" or "abnormal" NFC; the latter were further classified as having an "early", "active", "late", "normal" and "unclassifiable" Cutolo pattern (1). Each rater re-scored the NFC images of 10 selected patients (with NFC images rated by reader 1 as having good visibility at the first read).

The study was approved by the local ethics committee and all patients gave written informed consent.

STATISTICAL ANALYSIS

Semi-quantitative data were analyzed by calculating the mean score for capillary loss, disorganization of the microvascular array, giant capillaries, microhemorrhages and capillary ramifications. Two combinations were also calculated:

- giant capillaries + microhemorrhages = surrogate for vascular activity;
- disorganization + ramifications = surrogate for vascular damage.

Variables were analyzed for each finger and for all fingers by adding up the scores (separately for each parameter).

Inter-rater/intra-rater agreement was assessed by Cohen's kappa coefficients for qualitative variables and by intraclass correlation coefficients (ICC) for mean score values of abnormalities. Intra-rater agreement of the semi-quantitative assessment was analyzed using the Cronbach's alpha test.

Differences in scores between patients with/without digital ulcers (DUs) history and with/without Bosentan treatment were analyzed by logistic regression. All statistical analyses have been carried in IBM SPSS 20.0.

RESULTS

There were 512 images scored, of which 384 belonged to SSc patients. From these, 70 images of SSc patients and 13 of controls scored by rater 1, respectively 57 photos of patients and 10 of controls scored by rater 2 were considered to have low visibility. All images of the control subjects were rated as "normal". The demographic and clinical characteristics of the patients, including treatment, are displayed in Table 1.

TABLE 1. Demographic and clinical characteristics of the patients. Figures show number (%) unless otherwise specified

Variables, n (%)	All (n = 48)	dcSSc (n = 16)	lcSSc (n = 31)
Age, years, mean (SD)	55 (17)	44.5 (20)	59 (17)
Female	42 (87.5)	13 (81.3)	28 (90.3)
History of digital ulcers	29 (60.4)	12 (75.0)	17 (54.8)
Telangiectasiae	33 (66.4)	11 (68.8)	21 (67.7)
Autoantibodies			
Antinuclear antibodies	47 (97.9)	–	–
Anti-Scl70	21 (43.8)	12 (75.0)	9 (29.0)
Anti-centromere	9 (18.8)	9 (29.0)	9 (29.0)
Bosentan	11 (22.9)	5 (31.3)	6 (19.4)
Calcium channel blockers	28 (58.3)	7 (43.8)	21 (67.7)
Alprostadil	11 (22.9)	4 (25.0)	7 (22.6)
Cyclophosphamide	9 (18.8)	4 (25.0)	5 (16.1)
Methotrexate	11 (22.9)	4 (25.0)	7 (22.6)
Azathioprine	7 (14.6)	3 (18.8)	4 (12.9)
Corticosteroids	12 (25.0)	5 (31.3)	7 (22.7)

Assessment of abnormalities in a qualitative manner (normal/abnormal or Cutolo capillaroscopy patterns) showed moderate interrater agreement (k coefficient < 0.7). Results are displayed in Table 2 and Fig. 1.

TABLE 2. Qualitative assessment by identification of each capillaroscopic abnormality: inter-rater concordance and agreement by Cohen's kappa

Parameter	Concordance	Cohen's Kappa
Microhemorrhages	93%	0,847
Giant capillaries	94%	0,639
Ramified capillaries	87%	0,529
Disorganisation	85%	0,533
Decreased density*	90%	–

*Cohen's kappa for density couldn't be evaluated as one of the readers assesses all patients as having abnormal density.

For the semi-quantitative NFC assessment, inter-rater reliability ranged from good to excellent for mean score values of abnormalities in all examined fingers (ICC coefficients 0.745 to 0.897) and was excellent for activity (ICC coefficient of 0.923) and damage combinations (ICC coefficient of 0.918) – see Table 3. When scores were assessed in each finger, inter-rater reliability ranged also good to excellent for mean scores of abnormalities and also for the activity and damage combinations (ICC coefficients 0.781 to 0.867 for mean abnormalities scores and 0.713 to 0.856 for combinations), whereas for qualitative assessments inter-rater reliability was weaker (k coefficients < 0.7).

Intra-rater reliability was good to excellent for mean score values of abnormalities and activity and damage combinations in all fingers and in each finger for both raters (Table 4); for qualitative assessment, only one of the raters had good test-retest agreement (results not shown).

There was no difference in any NFC scores between patients with DUs history or Bosentan treatment (results not shown).

TABLE 3. Semi-quantitative assessment of each capillaroscopic abnormality (on a 0-3 scale) averaged for all 8 fingers examined: inter-reader reliability by intraclass correlation coefficient (ICC)

Parameter	ICC	95% confidence interval
Microhemorrhages	0,897	0,74 – 0,95
Giant capillaries	0,885	0,71 – 0,95
Ramified capillaries	0,804	0,49 – 0,92
Disorganisation	0,745	0,37 – 0,89
Density	0,866	0,65 – 0,94
Vascular activity score	0,923	0,81 – 0,97
Vascular destruction score	0,918	0,78 – 0,96

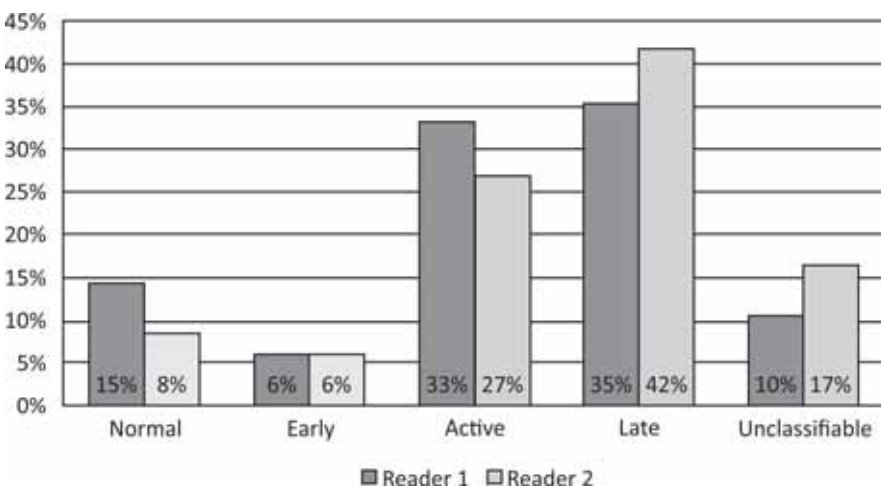


FIGURE 1. Qualitative assessment by Cutolo pattern identification: inter-reader agreement by Cohen's kappa was k = 0.514

TABLE 4. Semi-quantitative assessment of each capillaroscopic abnormality (on a 0-3 scale): intra-reader reliability by Cronbach's alpha

	Cronbach's alpha Reader 1	Cronbach's alpha Reader 2
Microhemorrhages	1,00	0,992
Giant capillaries	0,981	0,628
Ramified capillaries	0,976	0,986
Disorganisation	0,966	0,990
Density	0,996	0,933
Vascular activity score	0,990	0,837
Vascular destruction score	0,980	0,992

DISCUSSIONS

In this pilot study, we have shown good to excellent inter- and intra-rater reliability of semi-quantitative NFC assessment in patients with SSc, with indices suggesting superior reliability of this method over NFC qualitative assessment, in which both inter- and intra-rater agreement were only moderate. Among all semi-quantitative data, composite indices for microvascular disease activity and microvascular damage had the best reliability statistics, with both ICC and Cronbach's alpha > 0.9. These results are an important step towards developing semi-quantitative NFC as an outcome measure for the assessment of SSc-associated microvascular disease.

Our results are in concordance with those of Smith and coworkers, who were the first to use the semi-quantitative NFC method applied in our study and to assess its reliability (14,15). Similarly, the inter-rater agreement for qualitative assessment in their two-center, two-rater study was moderate to good, with Cohen's kappa indices ranging 0.5 to 0.7, with inter-rated concordance of 62-96% (weaker results for Cutolo patterns and stronger ones for the dichotomous classification of NFC into "normal" and "abnormal"). In semi-quantitative assessment, inter-rater reliability was good to excellent, with higher ICCs for capillary loss (0.96) and microhemorrhages (0.90) and lower ICCs for giant capillaries (0.86) and ramifications (0.64). Intra-rater agreement on qualitative assessment was superior in this study when compared to ours ($k=0.7-0.8$), but still lower than the intra-rater reliability of the semi-quantitative assessment (ICC=0.95 for every of the 4 abnormalities scored). The main difference between this study and ours relies in the simplification we made in recording NFC images: while Smith and coworkers recorded and analyzed 4 fields in each of

the examined fingers (using a videocapillaroscope with a magnification of x200), we used widefield stereomicroscopy with a magnification of x50, recording only one image per finger. This made our qualitative and semi-quantitative analyses easier and more rapid, while maintaining a satisfactory reliability. The main limitation in our study was the different experience in NFC of the readers (one of them being a medical student) and the relatively small cohort (although literature data shows that about 50 patients is a reasonable number for reliability studies) (16). Another limitation we have to acknowledge is the impossibility to use the same statistical method when evaluating reliability of the qualitative and the semi-quantitative assessment, respectively. While for nominal data the recommended method is Cohen's kappa, for numeric variables we had to use ICC and Cronbach's alpha; thus, comparing the results of qualitative and semi-quantitative, NFC has to take this aspect into account, although all statistical methods we have used are tests of reliability. While for Cohen's kappa moderate agreement is represented by values between 0.41-0.60 and good agreement by values between 0.61-0.80 (values > 0.80 suggesting almost perfect agreement), for ICC good agreement ranges between 0.60-0.74, all ICC values > 0.75 meaning excellent reliability (17,18).

In conclusion, this pilot study using widefield stereomicroscopy demonstrates very good intra- and inter-rater reliability of the semi-quantitative NFC assessment in a SSc cohort. The best results were achieved by the combined indices reflecting microvascular disease activity and microvascular damage, respectively. To demonstrate the utility of this method in clinical trials, longitudinal studies in larger cohorts are needed.

This study was performed as part of the project "Development of a computer-based nailfold videocapillaroscopy (NVC) system for longitudinal evaluation of patients with systemic sclerosis" (QUANTICAP), financed by the UEFIS-CDI PN-II-PT-PCCA-2013-4-1589 grant.

The results were presented as a poster at the ACR/AHRP 2016 Annual Meeting in Washington DC, November 11-16 (<http://acrabstracts.org/abstract/excellent-reliability-of-semiquantitative-nailfold-capillaroscopy-assessment-in-a-systemic-sclerosis-cohort-a-pilot-study/>).

REFERENCES

1. **Cutolo M., Pizzorni C., Sulli A.** Capillaroscopy. *Best Pract Res Clin Rheumatol* 2005; 19:437–52
2. **Maricq H.R., LeRoy E.C.** Patterns of finger capillary abnormalities in connective tissue disease by “wide-field” microscopy. *Arthritis Rheum* 1973; 16:619-28.
3. **van den Hoogen F., Khanna D., Fransen J., Johnson S.R., Baron M., Tyndall A., Matucci-Cerinic M., et al.** 2013 Classification Criteria for Systemic Sclerosis: An American College of Rheumatology/European League Against Rheumatism Collaborative Initiative. *Ann Rheum Dis* 2013; 72:1747-55.
4. Subcommittee for Scleroderma Criteria of the American Rheumatology Association Diagnostic and Therapeutic Criteria Committee. Preliminary criteria for the classification of systemic sclerosis (scleroderma). *Arthritis Rheum* 1980; 23:581-590.
5. **Gheorghiu A.M., Dobrota R., Predescu T., Grigorescu L., Soare A, Gorga M., Ionitescu R., et al.** Performance of the 2013 American College Of Rheumatology/European League Against Rheumatism Systemic Sclerosis Classification Criteria. *Ann Rheum Dis* 2016; 75(Suppl2): 410.
6. **Ingegnoli F., Boracchi P., Gualtierotti R., Lubatti C., et al.** Prognostic model based on nailfold capillaroscopy for identifying Raynaud’s phenomenon patients at high risk for the development of a scleroderma spectrum disorder: PRINCE (prognostic index for nailfold capillaroscopic examination). *Arthritis Rheum* 2008, 58: 2174-2182.
7. **Koenig M., Joyal F., Fritzler M.J., et al.** Autoantibodies and microvascular damage are independent predictive factors for the progression of Raynaud’s phenomenon to systemic sclerosis: a twenty-year prospective study of 586 patients, with validation of proposed criteria for early systemic sclerosis. *Arthritis Rheum.* 2008; 58: 3902-12.
8. **Cutolo M., Pizzorni C., Tuccio M., Burrioni A., Craviotto C., Basso M., et al.** Nailfold videocapillaroscopic patterns and serum autoantibodies in systemic sclerosis. *Rheumatology* (Oxford) 2004; 43:719-26
9. **Caramaschi P., Canestrini S., Martinelli N., et al.** Scleroderma patients nailfold videocapillaroscopic patterns are associated with disease subset and disease severity. *Rheumatology* (Oxford) 2007; 46:1566-9.
10. **Smith V., Decuman S., Sulli A., et al.** Do worsening scleroderma capillaroscopic patterns predict future severe organ involvement? a pilot study. *Ann Rheum Dis.* 2012 Oct; 71(10):1636-9.
11. **Aschwanden M., Daikeler T., Jaeger K.A., et al.** Rapid improvement of nailfold capillaroscopy after intense immunosuppression for systemic sclerosis and mixed connective tissue disease. *Ann Rheum Dis.* 2008 Jul; 67(7):1057-9.
12. **Miniati I., Guiducci S., Conforti M.I., et al.** Autologous stem cell transplantation improves microcirculation in systemic sclerosis. *Ann Rheum Dis.* 2009 Jan; 68(1):94-8.
13. **Cutolo M., Sulli A., Pizzorni C., Smith V.** Capillaroscopy as an Outcome Measure for Clinical Trials on the Peripheral Vasculopathy in SSc-Is It Useful? *Int J Rheumatol.* 2010; 2010. pii: 784947.
14. **Sulli A., Secchi M.E., Pizzorni C., Cutolo M.** Scoring the nailfold microvascular changes during the capillaroscopic analysis in systemic sclerosis patients. *Ann Rheum Dis.* 2008 Jun; 67(6):885-7.
15. **Smith V., Pizzorni C., De Keyser F., et al.** Reliability of the qualitative and semiquantitative nailfold videocapillaroscopy assessment in a systemic sclerosis cohort: a two-centre study. *Ann Rheum Dis.* 2010 Jun; 69(6):1092-6.
16. **DeVet HCW, Terwee C.B., Mokkink L.W., Knol D.L.** Reliability. In: *Measurement in Medicine.* Cambridge; 2011:96-154
17. **Landis J.R., Koch G.G.** The measurement of observer agreement for categorical data. *Biometrics* 1977; 33 (1): 159–174
18. **Cicchetti, Domenic V.** Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment* 1994. 6 (4): 284-290.