Reliability of nailfold capillary density measurement as a possible outcome measure for systemic sclerosis-related microangiopathy

Graham Dinsdale[1], Tonia Moore[2], Joanne Manning[2], Andrea Murray[1], Michael Berks[3], Philip Tresadern[3], Christopher Taylor[3], Neil O'Leary[4], Christopher Roberts[4], John Allen[5], Marina Anderson[6], Maurizio Cutolo[7], Roger Hesselstrand[8], Kevin Howell[9], Paula Pyrkotsch[6], Francesca Ravera[7], Vanessa Smith[10], Alberto Sulli[7], Marie Wildt[8], Ariane Herrick[1]

1. Centre For Musculoskeletal Research, University of Manchester, Manchester Academic Health Science Centre, Manchester, UK
2. Salford Royal Hospital NHS Foundation Trust, Salford, UK
3. Centre for Imaging Sciences, Institute of Population Health, University of Manchester, Manchester, UK
4. Centre for Biostatistics, Institute of Population Health, University of Manchester, Manchester, UK
5. Regional Medical Physics Department, Freeman Hospital, Newcastle upon Tyne, UK
6. Institute of Ageing and Chronic Disease, University of Liverpool, Liverpool, UK
7. Research Laboratory and Academic Division of Clinical Rheumatology, Dept. Internal Medicine, University of Genova, Italy
8. Department of Clinical Sciences, Section of Rheumatology, Lund University, Lund, Sweden
9. Institute of Immunity and Transplantation, University College London, Royal Free Campus, London, UK
10. Rheumatology, Ghent University Hospital, Ghent, Belgium

Background

Nailfold videocapillaroscopy allows non-invasive assessment of the microcirculation. Image annotation software allows tracking of changes over time; a potential outcome measure for systemic sclerosis (SSc) related microangiopathy. Our objective was to assess the reliability of capillary density measurement, known to be reduced in SSc[1].

Methods

124 patients (102 SSc, 22 primary Raynaud's phenomenon [PRP]), and 50 healthy controls (HC), underwent high-magnification (300x) videocapillaroscopy mosaic imaging of all 10 digits (1740 images).

Image subsets, sampled over disease categories, were randomly allocated to at least 2 of 10 capillaroscopy experts. These 'raters' used custom software to assess images. To assess intra-rater reliability, each rater performed repeat evaluations on an image subgroup. At least 6 images were assessed from each subject.

Raters marked distal vessel locations in an image. Vessel density was calculated as the total number of distal vessels divided by the Euclidean distance between the vessels at the horizontal extremities.

We examined: (1) the probability of raters marking sufficient (2 or more) distal vessels in an image (logistic mixed-effects model). Conditional on an image evaluation having sufficient distal vessels marked (2) distal vessel density (linear mixed-effects model).

Intra and inter-rater reliability was estimated with intra-class correlation coefficients from fitted model variance components.

Results

3463 images were evaluated. Each rater assessed a median (range) of 112 (87, 1406) unique images from 14 (9, 174) subjects. Same-rater repeat evaluations were performed on (median) 17% of images, and 904 images from 116 patients were evaluated by at least 2 raters.

Raters marked sufficient distal vessels in 79% of evaluations. Compared to HC, SSc and PRP images had odds ratios [95% CI] of sufficient distal vessels marked of 0.23 [0.14, 1.41] and 3.80 [0.37, 5.23] respectively. The mean vessel density in HC was 9.84 vessels/mm. Compared to HC, vessel density was lower in SSc (6.62) but not significantly different in PRP (9.58); respective differences [95% CI] were -3.22 [-3.88, -2.63] and -0.26 [-1.12, 0.63] vessels/mm.

Estimates of intra-rater reliability [95% CI] were 0.91 [0.89, 0.92] for vessel mark-up and 0.89 [0.87, 0.91] for vessel density. Corresponding estimates of inter-rater reliability were 0.51 [0.39, 0.76] and 0.56 [0.47, 0.64] respectively.

Conclusion

Mark-up rate differences between-groups are most likely due to differences in capillary architecture (capillary loss/damage in SSc patients). Density was unmeasurable in a sizable minority (21%) of image evaluations with potential implications for the representativeness of this measure. The high intra-(compared to inter-) rater reliability suggests that density could serve as outcome measure in prospective studies if the same rater examines images. Research on the impact of training on inter-rater reliability, and into more objective (automated) analysis methods is required to further develop this promising outcome measure.

1. Cutolo M et al. *Nature Rev Rheumatol* 2010;6:578-87