# Does nailfold capillaroscopy help predict future outcomes in systemic sclerosis? A systematic literature review

**Dolcie Paxton BSc [a,1] & John D Pauling BMedSci BMBS PhD FRCP [a,2,3]**

[1] University of Bristol Medical School, University of Bristol, Bristol, UK

[2] Department of Pharmacy and Pharmacology, University of Bath, Bath, UK

[3] Royal National Hospital for Rheumatic Diseases (at Royal United Hospitals), Bath, UK

[a] Dolcie Paxton & John D Pauling contributed equally to this paper as joint 1st authors

**Short title:** Prognostic value of NC in SSc

**Corresponding Author:**

Dr John D Pauling BMedSci PhD FRCP

Senior Lecturer & Consultant Rheumatologist

Department of Pharmacy and Pharmacology, University of Bath, Bath, UK

Royal National Hospital for Rheumatic Diseases, Upper Borough Walls, Bath, BA1 1RL

Tel: (0044) 1225 473 468 Fax: (0044) 1225 473 452 jdp32@bath.ac.uk

**Abstract**

**Background:** Nailfold capillaroscopy (NC) is an important diagnostic tool in systemic sclerosis (SSc). Confirmation of NC as a prognostic factor could facilitate earlier intervention and slow disease progression in SSc. We undertook a systematic literature review to evaluate the prognostic value of NC in predicting SSc disease progression

**Methods:** Standardised searches of EMBASE and MEDLINE were undertaken to identify longitudinal studies of adult subjects with SSc reporting the prognostic value of NC for any aspect of disease progression and/or survival. Non-English, non-original research, animal studies, non-adult studies and non-full length reports were excluded from the analysis (PROSPERO 2017:CRD42017071719). Wide heterogeneity in study design, prognostic factor measurement and study outcomes necessitated a qualitative data synthesis. The "QUality In Prognosis Studies" (QUIPS) risk-of-bias tool was used to assess study quality. Study selection, data extraction and risk-of-bias assessment were each undertaken independently by 2 reviewers and consensus reached where necessary.

**Results.** Of 942 retrieved articles, 18 studies fulfilled the inclusion criteria. The majority of studies (16/18, 89%) reported positive associations between baseline NC appearances (using a variety of qualitative, semi-quantitative and quantitative NC endpoints) and clinical outcomes including digital ulcer (DU) occurrence/healing, survival, disease progression (using domains of Medsger disease severity scale), calcinosis, skin progression, pulmonary arterial hypertension (PAH), and/or a composite analysis of "cardiovascular events". Application of the QUIPS tool identified a moderate-high risk of potential bias in 6/18 studies for study participation, 3/18 studies for study attrition, 10/18 for prognostic factor measurement, 5/18 for outcome measurement, 13/18 for confounders and 13/18 for statistical analyses. Study quality limited the strength of the conclusions drawn from these studies. The most important source of potential bias across the studies was insufficient adjustment for potential confounders; such as existing DU disease in studies evaluating future DU occurrence. Recent work suggests NC evolution is an important predictor of disease progression in SSc.

**Conclusions:** High levels of potential bias relating to study confounding and statistical analysis make it difficult to draw conclusions regarding the prognostic role of NC in SSc. There is strong evidence supporting an association between NC abnormalities (particularly capillary loss) and

disease severity (particularly vascular manifestations such as DU, calcinosis and PAH). Evolution of NC appearances may represent a more important predictor of disease progression which could have important implications for the future use of NC in the routine longitudinal assessment and management of SSc.

**Keywords:** Systemic Sclerosis, Scleroderma, Nailfold capillaroscopy, Prognosis, Clinical phenotype, Biomarkers, Systematic Literature Review

**Abbreviations**

Digital ulcers (DU)

International prospective register of systematic reviews (PROSPERO)

Nailfold capillaroscopy (NC)

Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA)

Pulmonary Arterial Hypertension (PAH)

Systemic sclerosis (SSc)

**Funding:** This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. Ms Dolcie Paxton undertook this work as part of a University of Bristol Medical School 3rd year Student Selected Components (SSC) placement.

**Conflicts of Interest:** None of the authors report any conflicts of interest relevant to the content of this work

**Introduction**

Systemic sclerosis (SSc) is a rare multisystem connective tissue disease characterised by a autoimmunity, aberrant tissue remodelling and vasculopathy [1]. The nailfold provides an opportunity to directly visualise the evolving obliterative microangiopathy of SSc and the characteristic nailfold capillary abnormalities have been well-described [2]. The morphological appearance of nailfold capillaries do not remain static and longitudinal studies have charted the progression of capillary abnormalities over time [3, 4]. The terminology used to classify nailfold capillary abnormalities reflect distinct patterns visualised at different stages of the disease with "early" changes characterised by the presence of few giant capillaries with few microhaemorrhages and "late" changes characterised by progressive capillary loss and disorganised neovascularisation [5] (Figure 1).

The diagnostic value of nailfold capillaroscopy (NC) in SSc [2, 6] has led to its inclusion in recent classification criteria for SSc [7, 8]. NC has also been shown to have an important predictive role identifying patients with Raynaud's phenomenon at risk of progression to overt SSc-spectrum disorders [9-11]. Another early observation was an association between severe NC abnormalities and internal organ involvement in SSc [12, 13]. Subsequent studies have identified associations between NC changes and a number of clinical features including cardiopulmonary disease [14], digital ulcers (DU) [15], calcinosis cutis [16], telangiectases [17], acro-osteolysis [16] and others [18]. A number of cross-sectional studies reporting associations between NC appearances and clinical phenotype have speculated about the prognostic and/or predictive value of NC in SSc but lacked longitudinal assessments to confirm this [14, 15, 18, 19].

Prognostic studies seek to predict future outcomes in individuals based on clinical and non-clinical factors [20]. Prognostic factors can offer insight into disease pathogenesis and support medical decision making [21]. If a prognostic role for NC were confirmed, it could be used to guide early intervention to modify the disease course of SSc [22]. We report the findings of a systematic literature review designed to evaluate the prognostic value of NC in predicting disease-specific outcomes in longitudinal studies of SSc.

## Methods and analysis

*Protocol development and review registration*

A study protocol was developed according to PRISMA-P guidelines [23] and registered in the International Prospective Register of Systematic Reviews (Registration: PROSPERO 2017:CRD42017071719).

*Eligibility criteria*

Studies reporting the relationship between NC and disease progression, prognosis and/or the presence/emergence of specific clinical features of systemic sclerosis were included in this systematic literature review. Applying the PICOS framework, we sought to evaluate publications that fulfilled the following study characteristics:

*Population:* Adults (18 years or older) with a diagnosis of systemic sclerosis.

*Intervention:* Studies reporting the use of NC as a prognostic factor. All methods of image acquisition and analysis were eligible for inclusion.

*Comparisons:* None

*Outcomes:* All relevant SSc-specific or SSc-associated outcomes including, but not limited to, organ-specific complications of SSc (e.g. digital ulcers, pulmonary arterial hypertension, scleroderma renal crisis etc.), clinical phenotype (e.g. SSc disease subsets), patient demographics (e.g. disease duration, gender, age), survival, health-related quality of life, functional status and other co-morbidities (e.g. concomitant cardiovascular disease).

*Study design:* Only longitudinal studies that reported outcomes at a time-point distinct from the baseline NC assessment (i.e. appropriate design to examine prognostic value) were eligible for data synthesis. Eligible study types included cohort studies (prospective or retrospective) and observational studies. The following studies were excluded from the analysis; animal studies, non-English language publications, studies of childhood/juvenile SSc,

studies of mixed patient populations (e.g. primary Raynaud's phenomenon, undifferentiated connective tissue diseases or overlap syndromes) in which a SSc cohort was not adequately reported, studies designed to develop/validate measurement scales, studies investigating the effectiveness of NC as a diagnostic/screening test, randomized controlled trials, case reports, qualitative research, non-original research publications (i.e., editorials, reviews), abbreviated reports (e.g. letters to editors) and conference proceedings.

*Information sources and search criteria*

Electronic searches were undertaken in Medline and EMBASE databases. The search criteria were developed in accordance with recommendations governing systematic reviews of evaluations of prognostic variables [24]. No publication date or language restrictions were applied to the searches. Full details of the specific search criteria applied within each database are available as supplementary material (Supplementary material 1).

*Study selection*

All titles and abstracts generated by the search were screened independently by two review authors (JP and DP) for relevance and eligibility of studies for full text review. Cohen's Kappa statistics were used to assess agreement between reviewers. Any discrepancies in agreement were reviewed together, with resolution through discussion, at each step of the study selection process. A grey search of potentially relevant cited articles identified during full text review was undertaken.

*Data extraction and synthesis from selected studies*

Data was independently extracted by both reviewers (JDP and DP) using a standardized form piloted during protocol development. The data extraction form collated relevant study details including date of publication, publication source, country of origin, study design, initial population of the study, study attrition, inclusion criteria, exclusion criteria, NC method, NC analysis, blinding of NC assessor, disease outcomes assessed (e.g. organ-specific manifestations of SSc), approach to statistical analysis and a summary of key findings. A qualitative data synthesis of the study findings was planned due to an anticipated high study heterogeneity in terms of study design, NC image acquisition/analysis and outcomes

rendering any meaningful attempt at meta-analysis inappropriate. There was an intention to contact corresponding authors where clarification was deemed necessary.

*Risk of Bias Assessment*

The quality of prognostic studies (and risk of bias) was independently assessed by 2 reviewers (JP and DP) using the QUality in Prognosis Studies (QUIPS) tool [25, 26] with discrepancies in agreement resolved through discussion where necessary (see Supplementary material 2 for further details).

**Results**

*Study selection*

Simultaneous searches in EMBASE (n=610) and Medline (n=332) were conducted on 12th July 2017 identifying a total of 942 articles. After removal of duplicate results (n=223), 719 articles were screened for eligibility in an initial title and abstract review resulting in the identification of 23 articles that proceeded to full text review. There was good agreement between reviewers at title and abstract screening with an unweighted Cohen's Kappa of 0.7 (see supplementary material 3 for details) with discordance in only 15/719 abstracts that was easily resolved through discussion, without the need for independent arbitration. Full text review of 23 manuscripts, led to further exclusion of 6 studies and the identification of 1 study deemed eligible for inclusion during "grey search" review of cited papers. This resulted in the identification of 18 papers that proceeded to full data extraction. The full study selection process is summarised in Figure 2. A summary of the main study attributes, study design and reported findings is presented in Table 1.

*Geographical participation and publication date*

The majority of the studies were undertaken in Europe (14/18, 78%) [27-40]; over half of which originated from 3 research teams. The remaining studies were led by centres in Canada [41, 42], Australia [43] and Brazil [44]. No prognostic studies of NC to date have been supported by US investigators. This is a developing field with the majority of studies published between 2013 and 2017 (10/18, 56%) with only 2 studies preceding 2009 [35, 42].

*Study characteristics*

The majority of studies were prospective longitudinal studies (13/18, 72%) with relatively short follow up. The duration of follow up in the prospective studies varied substantially according to the disease outcome of interest with 5 studies (38%) assessing outcomes within 6 months (all examining relationship between NC and DU occurrence/healing) [29-33], with the longest reported prospective follow up of 27 months (± 18 months) adopted in a study examining cardiovascular outcomes [39]. The follow up period reported in the retrospective studies was longer still (up to 20 years), particularly in the studies examining independent risk factors for survival [35, 42-44].

*Study populations*

A total of 3,385 SSc patients have been evaluated longitudinally in prognostic studies of NC. Multi-centre studies included the largest number of patients (total 2,445, with average of 272 per study) whereas total recruitment to single centre studies was lower (940 patients in 9 studies averaging 104 per study). The patients were recruited according to criteria used at the time of enrolment which included the 1980 preliminary ARA classification criteria [45] (applied in most studies whose publication pre-dated 2015), the LeRoy and Medsger classification criteria for early SSc [7] and latterly the 2013 ACR/EULAR classification criteria for SSc [8]. Some studies relied upon a clinician diagnosis of SSc [35, 38-42]. Most studies can be considered a representative SSc patient cohort and the QUIPS risk-of-bias assessment was considered "low" for the majority of studies (12/18, 66%, Table 2). Other studies incorporated eligibility criteria (e.g. the presence of specific NC abnormalities at baseline [30-33]) or *a priori* patient grouping to enrich populations with patients at a high risk of developing the outcome of interest [29] that limited the generalisability of the study findings. Other studies examined disease outcomes that necessitated a particular baseline clinical phenotype (e.g. the need for baseline DU when examining the relationship between NC abnormalities and DU healing [27]) or required the active exclusion of patients with relevant phenotype e.g. pre-existing cardiovascular risk factors [34, 39, 40]. Study attrition was generally reported to be low. Some studies did not report attrition despite the recruitment of large numbers of patients (e.g. n=219 in [30]) and/or a protracted follow-up period (e.g. 3 years in [34]). The QUIPS risk-of-

bias assessment was only considered “high” in a single study in which >1/3 of patients were lost to follow-up increasing the potential for selection bias [44].

*Prognostic factor measurement*

NC was one of several exploratory prognostic factors in each of the retrospective longitudinal studies [35, 41-43]. For the majority of prospective studies, NC was the principle or only PF measured although in others NC was one of a larger repertoire of exploratory PFs [39, 40]. NC image acquisition and analysis varied across studies. Where uniformity in approach existed, it was typically the result of studies arising from the same research group e.g. [30-33]. Most studies provided sufficient description of NC image acquisition and analysis but in others important details were missing [34, 35, 41], raising concerns as to whether NC had been measured using a standardised approach for all participants. The majority of studies reported NC assessment of all of 8 fingers with one study including the thumbs [44] whereas another restricted assessment to a single digit [43]. Various methods of NC image acquisition have been used at image magnifications ranging from 18x to 200x magnification (Table 1). One multicentre study used dermatoscopes rather than formal videocapillaroscopy to acquire NC images [41]. Seven studies reported blinding of the NC assessor (Table 1) although the extent to which PF assessment was entirely free of bias might be questioned. For example, the choice of “the most representative images” (one per finger) applied in some studies could be influenced by the presence of active DU or digital pitting at the time of the NC assessment.

A variety of qualitative, semi-quantitative and quantitative approaches to NC image analysis (sometimes in combination) have been applied in prognostic studies of NC and are summarised in Table 1. Commonly used methods include the qualitative NC classification approach initially proposed by Cutolo et al. [5] which was included in 7/18 (39%) of studies from a variety of centres, and the Capillaroscopic Skin Ulcer Risk Index (CSURI) used in 4 studies (22%, all originating from the same group [30-33]). Across all studies, the bias rating for PF measurement using the QUIPS tool was considered “moderate” for 9/18 (50%) of the studies and “high” for 1 study which failed to report the method of NC image acquisition (Table 2). More recent prospective studies have reported a standardised approach to NC

image acquisition and analysis that would be more easily reproducible and less amenable to bias [28, 29].

*Outcome measurement*

The studies reported associations between baseline NC appearances (using a variety of qualitative, semi-quantitative and quantitative NC endpoints) and clinical outcomes including digital ulcer (DU) occurrence/healing [26, 28-33, 35], survival [34, 41-43], disease severity progression (including peripheral vascular using domains of disease severity scale [DSS]) [27, 36, 37], calcinosis [40], skin progression [27], pulmonary arterial hypertension [27, 39], and/or composite analysis of "cardiovascular events" [38]. Outcome measure assessments generally conformed to accepted definitions/validated instruments although some studies did not provide detailed descriptions of relevant clinical variable definitions (e.g. what constitutes the presence/healing of DU). The approach to outcome measure assessment was generally unbiased and rated "low" for most studies (13/18, 72%) using the QUIPS tool. There were examples in which the assessment of the clinical outcome (e.g. whether there had been the occurrence of new DU and/or healing of existing DU) could have been influenced by the baseline NC findings (Table 2).

*Reported prognostic associations of NC in SSc*

The majority of studies (17/18, 94%) reported at least 1 positive association between NC findings and clinical outcomes raising the possibility of reporting bias (Table 1). Across all of the studies, it was nailfold capillary loss (a feature considered to represent "late" disease and often classified as such) assessed using a variety of quantitative, semi-quantitative and qualitative approaches that was most often reported as being a predictor of a broad range of future SSc outcomes ranging from DU healing [27, 33], new DU occurrence [29-34], progression of peripheral vascular disease severity according to the DSS [36-38], progression of lung disease severity according to DSS [37], mortality [43, 44], cardiovascular events [39], pulmonary arterial hypertension [PAH] [40] and calcinosis [41]. Interpretation of previous studies is limited by moderate-to-high risk of bias across each of the domains of the QUIPS tool making it difficult to draw strong conclusions regarding the prognostic value of NC in SSc. A "high" risk-of-bias rating for at least 1 of the QUIPS domains was present in 12/18 (67%)

studies. The main issues with study quality were related to confounding and statistical analysis with 12/18 (67%) of studies rated as “high” across one or both of these domains (see Table 2 and the more detailed discussion that follows this section).

The most recently published study examining the prognostic role of NC in SSc is notable for the quality of the study design (and was the only study rated “low” risk-of-bias across all domains of QUIPS) and for the fact that it challenges the findings of earlier work [28]. This study reported an independent association between the presence of neoangiogenesis (another feature of “late” disease) on baseline NC and future lung vascular progression [28]. Baseline capillary drop out was *not* an independent risk factor for disease progression on multivariate analysis but there were novel associations between NC evolution (assessed annually for 3 years) and disease progression with progressive capillary loss being associated with a broad range of worsening disease outcomes (Table 1) [28]. An increased number of giant capillaries, meanwhile, appears to confer a protective effect against new DU and overall disease progression suggesting effective vascular remodelling can occur in SSc [28].

*Study confounding*

There was a high likelihood that the observed effect of NC on the outcome of interest could have been distorted by other factors related to both NC and outcome in many of the reported studies. Overall, 11/18 (61%) studies were rated as “high” risk of bias due to inadequate accounting for potential confounders (Table 2). For example, some studies did not supply baseline clinical data to confirm whether the clinical phenotype had evolved during the intervening period between NC and outcome assessment [37]. Inadequate adjustment for confounding might explain conflicting accounts as to whether NC abnormalities are a risk factor for mortality [35, 42-44]. Kayser et al. reported high avascular scores on NC as an independent predictor of mortality in a study that included 12 deaths, 4 of which were attributable to pulmonary arterial hypertension (PAH) [44]. At the time of NC assessment, 4 patients had echocardiographic evidence of PAH (with differences in the proportion with PAH at baseline in the “deceased” cohort versus the “alive” cohort approaching statistical significance [36% vs. 13%, p=0.07]) [44]. The presence of PAH, however, was excluded from the multivariate cox proportional hazards model reported to demonstrate the association

between the avascular score and mortality [44]. Similarly, Hissaria et al. reported a significant hazard regression ratio of 1.60 (95%CI 1.06-2.42) for mortality using categorical NC patterns treated as 0-3 continuous variables (grade 3 indicative of capillary drop out) but it was unclear whether this analysis was adjusted for the presence of diffuse scleroderma, interstitial lung disease (ILD) and the presence of anti-Scl-70 autoantibodies; each of which had been found to be independently associated with *both* capillary drop out and mortality in separate analyses reported in the study [43]. Similarly, studies reporting a predictive role of NC abnormalities in predicting future cardiovascular events and PAH did not incorporate important potential confounders (such as baseline echocardiographic parameters and serum brain natriuretic protein) in multivariate logistic regression analyses despite observed baseline differences in these variables between patients with/without future cardiovascular events/PAH [39, 40].

Inadequate accounting for confounders may have influenced studies reporting the prognostic value of NC in predicting future DU occurrence (e.g. [30-34, 36]). A recent well-designed study (considered low risk of bias for study confounding) examining the association between NC abnormalities and future DU identifiedclinical features such as the presence and number of active DU at baseline as the strongest predictors of new DU occurrence; although capillary drop out was an independent predictor of new DU within the cohort of patients with a history of DU [29].

The impact of adequately adjusting for potential confounders is highlighted by the findings of the aforementioned recent prospective study whose findings conflict with previous studies examining similar outcomes [28]. Univariate analysis identified severe loss of capillaries at baseline NC as a predictor of overall disease progression, the occurrence of new DU, lung vascular progression, skin progression and worsening of the DSS, however, none of these associations persisted on multivariate analysis [28]. In contrast, neoangiogenesis at baseline NC retained a significant association with the lung vascular progression (HR 11.12, 95%CI 1.19-103.79, p=0.036) with a strong trend for overall disease progression (HR 2.77, 95% CI 0.89-8.56) on multivariate analysis that controlled for relevant confounders [28].

*Statistical analysis and reporting*

A number of studies applied analytical methods that limit interpretation of the study findings. For example, the use of Mann Whitney U to compare the distribution of quantitative NC findings at baseline between patients whose DUs healed versus those whose DUs persisted (capillary density 4.9 vs. 3.3, p=0.05) supports an association between baseline NC changes and DU healing but is unable to examine the independent prognostic value of NC [27]. Similarly, a number of studies have applied receiver operating curve characteristics to confirm an association between baseline NC abnormalities and future digital ischaemic outcomes [31-33, 36]. This approach does not allow for adjustment for important baseline confounders (such as the presence of DU at baseline) and therefore limits the conclusions drawn concerning the independent prognostic value of NC. Once again, adjustment for confounding is paramount and multivariate cox proportional hazards regression modelling provides an opportunity to achieve this; although this approach is not beyond reproach if relevant confounding factors are not incorporated into the model (as highlighted previously [43]). Moreover, regression analysis models struggle to manage situations in which a confounding factor (such as history of DU) completely separates a predictor variable (such as NC features) for a given outcome (new DU occurrence). The statistical phenomena of quasi-separation may explain the exceptionally high odds ratio and broad confidence intervals (OR 99.95, 95%CI 20.03-498.86) obtained following multivariate analysis in the study examining the prognostic value of the CSURI in predicting new DU [30, 46]. The strength of the association between NC abnormalities and future DU was certainly considerably greater than in subsequent larger longer-term multicentre initiatives [29]. A propensity for smaller studies of single prognostic marker studies to sometimes report higher relative risks than larger studies has been noted previously [47]. Overall, 13/18 (72%) of studies were rated as moderate/high risk of bias related to analysis or reporting.

**Discussion**

To our knowledge, this is the first study to systematically review and critically appraise studies reporting the prognostic value of NC in SSc. This is a rapidly developing field with almost half of all prognostic studies published within the last 3 years. We identified 18 studies examining

the prognostic value of NC in SSc; the majority of which reported a prognostic role for NC abnormalities in predicting a broad range of disease-related outcomes, including mortality. Issues with study quality (relating primarily to study confounding and statistical analysis) limits the strength of conclusions that can be drawn from many of the studies reviewed and the evidence confirming a prognostic role for single time-point NC assessments in predicting disease outcomes in SSc at the present time is limited. Many of the studies identified in this review were early exploratory studies that would be termed "Phase I" prognostic factor studies which are known to be vulnerable to type one (false positive) errors [21]. Common to observations in other fields of prognostic study research [48], many of the identified studies focused on the importance of NC (a relatively high-cost biomarker), whilst over-looking the potential relative value of basic clinical assessments (such as history/presence of active DU); potentially over-estimating the added value gained from NC in predicting future SSc outcomes. A number of studies examined multiple prognostic factors and outcomes raising the potential for selective reporting bias in studies lacking a robust *a priori* statistical analysis plan [48]. Similar issues have been noted during efforts to synthesise evidence from prognostic studies in other disease areas [47]. Standards for the design, analysis and reporting of prognostic studies have been proposed that could support future data synthesis efforts in this field [21, 47-50].

The studies identified in this review do contribute to a larger body of evidence from cross-sectional studies identifying strong *associations* between NC abnormalities (namely capillary drop out) and a broad range of disease-related outcomes, particularly concerning vascular features such as DU disease and PAH [12-19]. Whilst the association between NC abnormalities and clinical features sheds important light on the pathophysiology of SSc, it does little to support treatment decisions in clinical practice (and it is possible such NC abnormalities develop after the outcome of interest). Confirming a prognostic value of NC could therefore facilitate earlier intervention and support improved outcomes. The recent work by Avouac et al. [28] is of major importance in this regard for 3 reasons. Firstly, the design, conduct, analysis and reporting of this large prospective observational cohort study were largely free of the aforementioned biases that can influence the interpretation of prognostic studies (it was the only study to be rated "low" across all of the domains of the

QUIPS tool). Secondly, the strength of the prognostic value of *baseline* NC was weaker than in other studies and only partially confirmed after multivariate Cox analysis. Thirdly, this study identified an important relationship between the evolution of NC abnormalities and future disease progression in SSc. Of particular interest were changes in NC appearances that appeared protective (increased number of giant capillaries) against SSc outcomes, whereas increased capillary loss portended worse outcomes across a broad range of clinical outcomes. These findings, if replicated, have important ramifications for the routine assessment of SSc and strengthens the case for serial NC to be incorporated into routine clinical monitoring of SSc disease progression.

This work has identified a number of factors that should be considered for future prognostic studies of NC in SSc. Common to many human diseases, the disease progression in SSc is likely influenced by patient demographics, clinical phenotype, and a complex interplay of a wide range of serological, genetic, environmental, treatment and psychosocial factors. As such, a multi-variable approach to the design and analysis of future prognostic studies of SSc is essential if we are to confirm a specific role for NC in the routine monitoring of SSc [20]. Ensuring sufficient observer blinding during measurement of the both NC (or other prognostic factor) and outcomes is another important aspect to consider in future prospective cohort prognostic studies [20]. It has been suggested that prognostic research of complex disease requires several hundred outcome events to avoid overestimations of the predictive performance of the model [20]. It has also been suggested that the analysis and reporting of prognostic studies "should focus on absolute risk estimates of outcomes given combinations of predictor values" [20]. Regression modelling with stepwise variable selection (e.g. Cox proportional hazards regression model that was used in many of the identified studies) sometimes performs badly in validation cohorts and cannot be confidently applied unless outcome events are several fold higher than the number of potential prognostic variables studied [21]. Moreover, use of the same data to develop the model and estimate regression coefficients can lead to biased estimates whose statistical significance can be misleading [21]. Such issues surrounding sample size present a particular problem for prognostic research in rare diseases such as SSc.

There were a number of limitations to the present work. The exclusion of unpublished studies (e.g. conference abstracts) and abbreviated reports from this review increases the risk of publication and reporting bias and may have contributed to the high number of positive studies. We took this decision to ensure we had sufficient information available to fully assess each of the studies subject to full data extraction. We did identify one example of a letter to editor [51] reporting the abridged findings of a longitudinal study examining the prognostic value of NC in SSc that was not eligible for inclusion in our study analysis due to our study eligibility criteria. In contrast to clinical trials and systematic literature reviews, there is currently no agreed standard for registering prognostic study protocols to facilitate easier identification of unpublished work. Relatively small positive studies (many undertaken without peer-reviewed funding) have a tendency to bias the literature of prognostic research [48]. We don't believe this has biased our principal conclusions as our critical appraisal of the quality of published studies has led us to caution against over-estimating the prognostic value of NC in SSc. We have raised a number of methodological issues that should be considered in the design of future prognostic studies in this field.

## Conclusions

The quality of many of the previous prognostic studies of NC in SSc have limited the conclusions that can be drawn from such work. More recent high-quality studies have reported findings which could have major implications for the positioning of NC in the longitudinal assessment and management of SSc in the future.

## Acknowledgements

The authors wish to thank Dr Victoria Flower for obtaining and supplying the example NC images presented in Figure 1.

## References

1. Denton, C.P. and D. Khanna, *Systemic sclerosis.* Lancet, 2017.

2. Maricq, H.R. and E.C. LeRoy, *Patterns of finger capillary abnormalities in connective tissue disease by "wide-field" microscopy.* Arthritis Rheum, 1973. **16**(5): p. 619-28.
3. Sulli, A., et al., *Progression of nailfold microvascular damage and antinuclear antibody pattern in systemic sclerosis.* J Rheumatol, 2013. **40**(5): p. 634-9.
4. Sulli, A., et al., *Timing of transition between capillaroscopic patterns in systemic sclerosis.* Arthritis Rheum, 2012. **64**(3): p. 821-5.
5. Cutolo, M., et al., *Nailfold videocapillaroscopy assessment of microvascular damage in systemic sclerosis.* J Rheumatol, 2000. **27**(1): p. 155-60.
6. Maricq, H.R., et al., *Diagnostic potential of in vivo capillary microscopy in scleroderma and related disorders.* Arthritis Rheum, 1980. **23**(2): p. 183-9.
7. LeRoy, E.C. and T.A. Medsger, Jr., *Criteria for the classification of early systemic sclerosis.* J Rheumatol, 2001. **28**(7): p. 1573-6.
8. van den Hoogen, F., et al., *2013 classification criteria for systemic sclerosis: an American College of Rheumatology/European League against Rheumatism collaborative initiative.* Arthritis Rheum, 2013. **65**(11): p. 2737-47.
9. Maricq, H.R., et al., *Microvascular abnormalities as possible predictors of disease subsets in Raynaud phenomenon and early connective tissue disease.* Clin Exp Rheumatol, 1983. **1**(3): p. 195-205.
10. Maricq, H.R., A.B. Weinberger, and E.C. LeRoy, *Early detection of scleroderma-spectrum disorders by in vivo capillary microscopy: a prospective study of patients with Raynaud's phenomenon.* J Rheumatol, 1982. **9**(2): p. 289-91.
11. Koenig, M., 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**(12): p. 3902-12.
12. Chen, Z.Y., et al., *Association between fluorescent antinuclear antibodies, capillary patterns, and clinical features in scleroderma spectrum disorders.* Am J Med, 1984. **77**(5): p. 812-22.
13. Maricq, H.R., G. Spencer-Green, and E.C. LeRoy, *Skin capillary abnormalities as indicators of organ involvement in scleroderma (systemic sclerosis), Raynaud's syndrome and dermatomyositis.* Am J Med, 1976. **61**(6): p. 862-70.
14. Markusse, I.M., et al., *Predicting cardiopulmonary involvement in patients with systemic sclerosis: complementary value of nailfold videocapillaroscopy patterns and disease-specific autoantibodies.* Rheumatology (Oxford), 2017. **56**(7): p. 1081-1088.
15. Ghizzoni, C., et al., *Prevalence and evolution of scleroderma pattern at nailfold videocapillaroscopy in systemic sclerosis patients: Clinical and prognostic implications.* Microvasc Res, 2015. **99**: p. 92-5.
16. Morardet, L., et al., *Late Nailfold Videocapillaroscopy Pattern Associated With Hand Calcinosis and Acro-Osteolysis in Systemic Sclerosis.* Arthritis Care Res (Hoboken), 2016. **68**(3): p. 366-73.
17. Hurabielle, C., et al., *Skin Telangiectasia and the Identification of a Subset of Systemic Sclerosis Patients With Severe Vascular Disease.* Arthritis Care Res (Hoboken), 2016. **68**(7): p. 1021-7.
18. Ingegnoli, F., et al., *Nailfold capillaroscopy in systemic sclerosis: data from the EULAR scleroderma trials and research (EUSTAR) database.* Microvasc Res, 2013. **89**: p. 122-8.

19. Riccieri, V., 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**(6): p. 992-7.
20. Moons, K.G., et al., *Prognosis and prognostic research: what, why, and how?* BMJ, 2009. **338**: p. b375.
21. Simon, R. and D.G. Altman, *Statistical aspects of prognostic factor studies in oncology.* Br J Cancer, 1994. **69**(6): p. 979-85.
22. Pauling, J.D., *Could nailfold videocapillaroscopy usher in a new era of preventative disease-modifying therapeutic intervention in systemic sclerosis?* Rheumatology (Oxford), 2017. **56**(7): p. 1053-1055.
23. Shamseer, L., et al., *Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: elaboration and explanation.* Bmj, 2015. **349**: p. g7647.
24. Altman, D.G., *Systematic reviews of evaluations of prognostic variables.* BMJ, 2001. **323**(7306): p. 224-8.
25. Hayden, J.A., et al., *Assessing bias in studies of prognostic factors.* Ann Intern Med, 2013. **158**(4): p. 280-6.
26. Hayden, J.A., P. Cote, and C. Bombardier, *Evaluation of the quality of prognosis studies in systematic reviews.* Ann Intern Med, 2006. **144**(6): p. 427-37.
27. Alivernini, S., et al., *Skin ulcers in systemic sclerosis: determinants of presence and predictive factors of healing.* J Am Acad Dermatol, 2009. **60**(3): p. 426-35.
28. Avouac, J., et al., *Sequential nailfold videocapillaroscopy examinations have responsiveness to detect organ progression in systemic sclerosis.* Seminars in Arthritis and Rheumatism, 2017. **47**(1): p. 86-94.
29. Cutolo, M., et al., *Nailfold Videocapillaroscopic Features and Other Clinical Risk Factors for Digital Ulcers in Systemic Sclerosis: A Multicenter, Prospective Cohort Study.* Arthritis & Rheumatology, 2016. **68**(10): p. 2527-2539.
30. Manfredi, A., et al., *Prediction risk chart for scleroderma digital ulcers: A composite predictive model based on capillaroscopic, demographic and clinico-serological parameters.* Clinical Hemorheology and Microcirculation, 2015. **59**(2): p. 133-143.
31. Sebastiani, 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**(5): p. 688-94.
32. Sebastiani, M., et al., *Capillaroscopic Skin Ulcers Risk Index (CSURI) calculated with different videocapillaroscopy devices: how its predictive values change.* Clin Exp Rheumatol, 2013. **31**(2 Suppl 76): p. 115-7.
33. Sebastiani, M., et al., *Predictive role of capillaroscopic skin ulcer risk index in systemic sclerosis: a multicentre validation study.* Ann Rheum Dis, 2012. **71**(1): p. 67-70.
34. Silva, I., et al., *Endothelial Dysfunction and Nailfold Videocapillaroscopy Pattern as Predictors of Digital Ulcers in Systemic Sclerosis: a Cohort Study and Review of the Literature.* Clinical Reviews in Allergy & Immunology, 2015. **49**(2): p. 240-252.
35. Simeon, C.P., et al., *Mortality and prognostic factors in Spanish patients with systemic sclerosis.* Rheumatology (Oxford), 2003. **42**(1): p. 71-5.
36. Smith, V., et al., *Nailfold capillaroscopy for day-to-day clinical use: construction of a simple scoring modality as a clinical prognostic index for digital trophic lesions.* Ann Rheum Dis, 2011. **70**(1): p. 180-3.

37. Smith, V., et al., *Do worsening scleroderma capillaroscopic patterns predict future severe organ involvement? a pilot study.* Ann Rheum Dis, 2012. **71**(10): p. 1636-9.
38. Smith, V., et al., *Nailfold Capillaroscopy for Prediction of Novel Future Severe Organ Involvement in Systemic Sclerosis.* Journal of Rheumatology, 2013. **40**(12): p. 2023-2028.
39. Voilliot, D., et al., *Cardiovascular outcome in systemic sclerosis.* Acta Cardiol, 2015. **70**(5): p. 554-63.
40. Voilliot, D., et al., *Prediction of new onset of resting pulmonary arterial hypertension in systemic sclerosis.* Archives of Cardiovascular Diseases, 2016. **109**(4): p. 268-277.
41. Baron, M., et al., *Calcinosis is associated with digital ischaemia in systemic sclerosis-a longitudinal study.* Rheumatology, 2016. **55**(12): p. 2148-2155.
42. Scussel-Lonzetti, L., et al., *Predicting mortality in systemic sclerosis: analysis of a cohort of 309 French Canadian patients with emphasis on features at diagnosis as predictive factors for survival.* Medicine (Baltimore), 2002. **81**(2): p. 154-67.
43. Hissaria, P., et al., *Survival in scleroderma: results from the population-based South Australian Register.* Intern Med J, 2011. **41**(5): p. 381-90.
44. Kayser, C., et al., *Nailfold capillaroscopy abnormalities as predictors of mortality in patients with systemic sclerosis.* Clin Exp Rheumatol, 2013. **31**(2 Suppl 76): p. 103-8.
45. *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**(5): p. 581-90.
46. Albert, A. and J.A. Anderson, *On the Existence of Maximum-Likelihood Estimates in Logistic-Regression Models.* Biometrika, 1984. **71**(1): p. 1-10.
47. Hemingway, H., et al., *Evaluating the quality of research into a single prognostic biomarker: a systematic review and meta-analysis of 83 studies of C-reactive protein in stable coronary artery disease.* PLoS Med, 2010. **7**(6): p. e1000286.
48. Hemingway, H., R.D. Riley, and D.G. Altman, *Ten steps towards improving prognosis research.* BMJ, 2009. **339**: p. b4184.
49. McShane, L.M., et al., *REporting recommendations for tumor MARKer prognostic studies (REMARK).* Nat Clin Pract Urol, 2005. **2**(8): p. 416-22.
50. Riley, R.D., et al., *Prognosis Research Strategy (PROGRESS) 2: prognostic factor research.* PLoS Med, 2013. **10**(2): p. e1001380.
51. Sebastiani, M., et al., *Validation study of predictive value of capillaroscopic skin ulcer risk index (CSURI) in scleroderma patients treated with bosentan.* Clin Exp Rheumatol, 2015. **33**(4 Suppl 91): p. S196.

## Tables and Figures

## Figure 1. Examples of capillary appearances at the nailfold of a healthy control and in systemic sclerosis

**A) Appearances at 50x magnification in a healthy control demonstrating uniformly-spaced nailfold capillaries of similar size and length. The distal row typically has a capillary density of >10 capillaries per mm (1mm size bar presented for interpretation); B) Appearances at 200x magnification of the "early" changes of systemic sclerosis [5] demonstrating few giant capillaries (no haemorrhages visible in this image) ; C) Appearances at 200x magnification of the "active" changes of systemic sclerosis [5] demonstrating the presence of giant capillaries, some reduction in capillary density and the remnants of a microhaemorrhage within the cuticle; D) The characteristic appearances at 200x magnification of the "late" changes of systemic sclerosis [5] demonstrating significant capillary loss and extensive disorganised neoangiogenesis.**

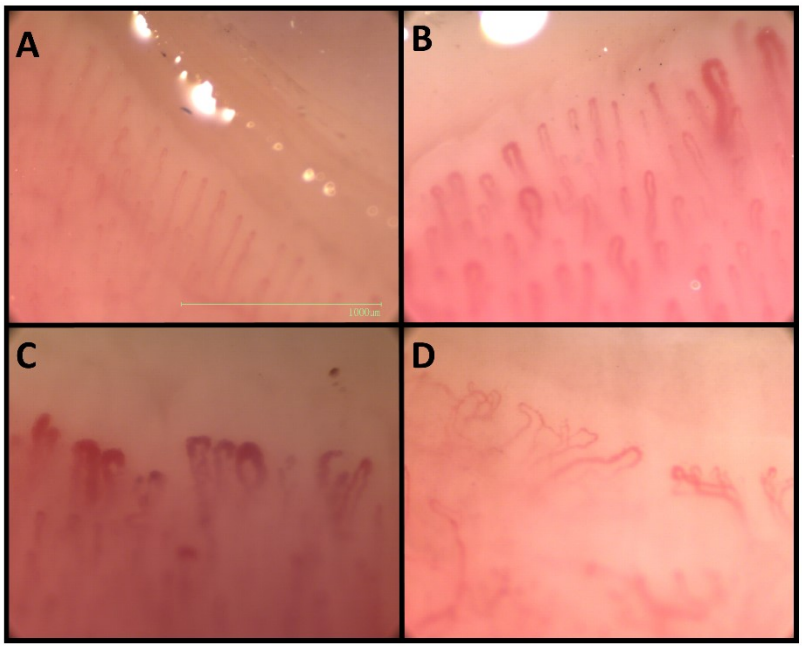

**Figure 2. Study selection flow diagram presented according to PRISMA statement**
NC, nailfold capillaroscopy

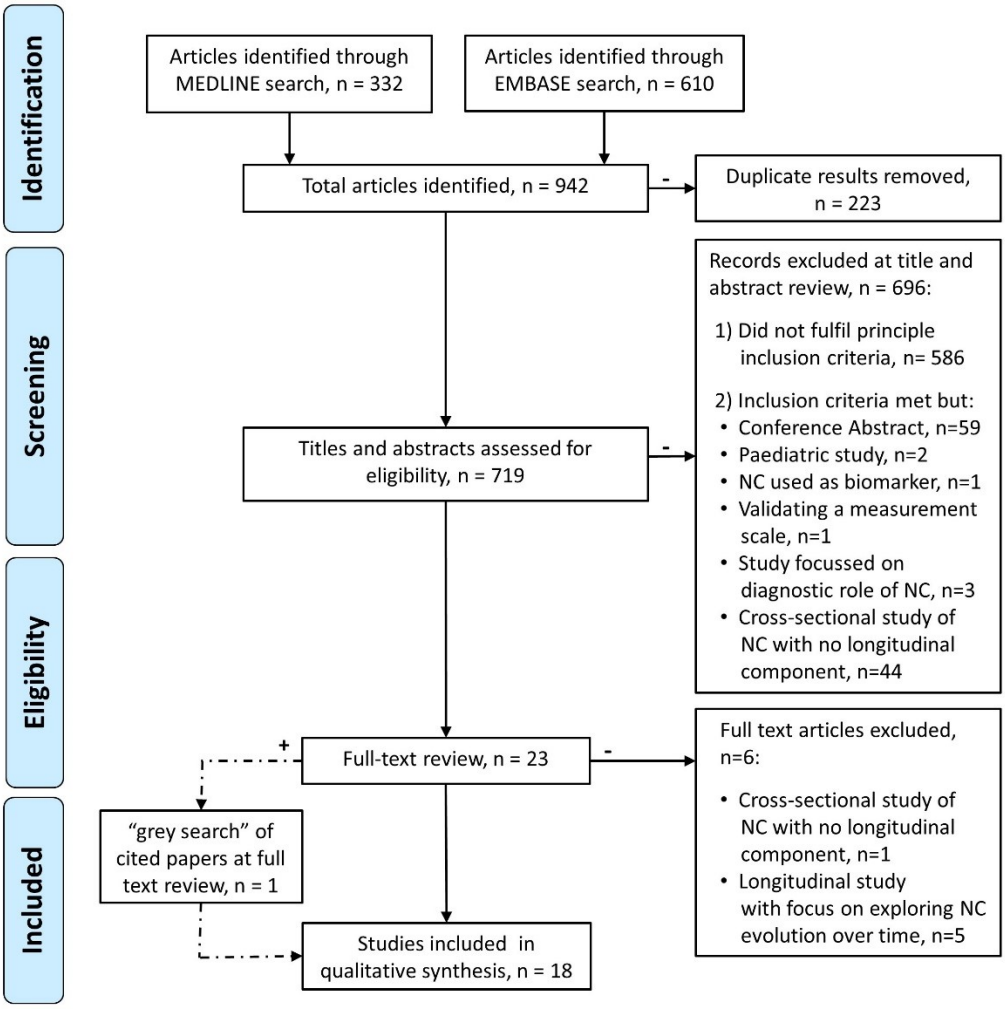

**Table 1. Summary table of study characteristics and major findings of studies examining the prognostic value of nailfold capillaroscopy in systemic sclerosis**

| Author & Date | Origin | Design | Description | Study Population | Attrition | Classification Criteria for SSc | Exclusion Criteria | NC Method | NC Analysis | Assessor blinding | Disease outcome(s) | Statistical methods to investigate relationship between PF and outcome | Summary of key reported findings and additional comments |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Scussel-Lonzetti et al. 2002 [42] | Canada | R L | Risk factors for mortality in SSc | 309 SSc pts (1st seen 1984 to 1999) | NC only available for 294 | Clinician diagnosis of SSc | Nil | All digits of both hands. 8-50X magn. | Categorical analysis of: (i) Cap dilatation (0-4 scoring) (ii) Avascular areas (scored A-D). | Yes | Mortality (census 2000) | Univariate Cox proportional hazards regression analysis followed by multivariate analysis using stepwise Cox regression analysis examining multiple variables of interest (including NC). | 66 patients died during follow-up. Capillary loss and capillary dilatations at baseline assessment not independent risk factors for mortality. Follow-up period for each patient at time of mortality assessment not presented (disease duration varied from <1 year to 17.5 years). Unclear whether potential confounding effect of disease duration included in analysis. |
| Simeón et al. 2003 [35] | Spain | R L | Risk factors for mortality in SSc | 79 SSc pts (1st seen 1976 to 1996) | NC only available for 69 | Clinician diagnosis SSc and disease duration <15 yrs; 65/79 fulfilled 1980 ARA criteria | Nil | 2nd-5th digits of both hands (magn. not reported) | Qual assessment of cap appearances into "active" (predominance of cap loss) and "slow" (megacapillaries and no cap loss) | Not reported | Mortality | Mortality rate expressed as person-years. The log-likelihood ratio x2 test was used as part of a crude analysis examining factors associated with survival. The Cox proportional hazards model was used to identify independent prognostic factors (including NC). | There were 12 deaths (11 of SSc-organ complications and 1 of cancer). On crude analysis, lung involvement (FVC <70%), PAH, SSc renal crisis, an "active" capillaroscopic pattern (p=0.02), pericardial effusion and age >60yrs at diagnosis were associated with shorter survival. On multivariate analysis, only age >60yrs at diagnosis, FVC <70% and SSc renal crisis were independent prognostic factors suggesting the initial observed effect of NC on survival on crude analysis distorted by confounders. |
| Alivernini et al. 2009 [27] | Italy | P L | Prognostic factors related to DU healing at 20 months. | 130 in initial study (34 baseline DU eligible for longitudinal analysis) | 4 with baseline DU not included in analysis. | 1980 ARA classification criteria for SSc | Nil | Not reported | Quant analysis of mean value of megacapillaries and cap density. Qual analysis of avascular areas, micro-haemorrhages & cap morphology. | Not reported | DU healing in pts with DU at baseline at 20-months follow-up. | Statistical methods applied to assessing the relationship between NC and DU healing not reported. The method section suggests Mann Whitney U used to compare means between groups. | 34 pts had DU at baseline. DUs healed in 11/30 vs. 19/30 persisting at 20 months. Patients whose DU healed had a higher capillary density on baseline NC (4.9±2.3) compared with those with persistent DU (3.3±1.3, p=0.05). NC magnification, number of digits assessed, number of images obtained and analysed not reported. Incomplete description of NC image analysis. Definition of DU healing not reported. Statistical methods applied to examining prognostic value of NC not reported. |
| Sebastiani et al. 2009 [31] | Italy | P L | Develop and validate a novel NC scoring method for predicting new DU at 3 months. | 120 pts. 35 developed new DU at 3 months. | None reported | 1980 ARA classification criteria for SSc | Nil | 2nd-5th digits of both hands. 200x magn. | Single image from each digit used to quantify lowest number of capillaries (N), the highest number of megacapillaries (M) and max diameter of largest megacapillary (D). | Yes | Occurrence of new DU at 3-month follow-up | ROC curve analysis examined optimal cut-offs for individual NC parameters (including N, M and D) comparing those with/without new DU. The composite capillaroscopic skin ulcer risk index (CSURI) was devised ($MxD/N^2$). ROC curve analysis used to determine optimal CSURI cut-off for "predicting" new DU at 3-months. | ROC curve analysis (comparing patients with new DU at 3 months versus those without) revealed an optimal CSURI cut off of 2.94 with PPV of 73.3%. No baseline assessment reported to confirm DU at 3 months were new rather than refractory. High rate of new DUs at 3 months (29.2%) suggests some might have been present at baseline. Description of model building strategy for development of CSURI not reported. ROC curve analysis cannot account for potential confounders (such as presence of DU at baseline) which might have biased choice of NC image for analysis). |
| Hissaria et al. 2011 [43] | Australia | R L | Risk factors for mortality in SSc | NC in 127 pts (single centre) from registry 786 pts (1st seen 1993-2007) | 24 lost to follow up | Modified 1980 ARA classification criteria for SSc | Nil | 4th digit of both hands. 16-18x magn. | Cap dilatation and drop-out categorised as normal, mild, moderate and severe. NC changes scored as (0) normal, (1) dilatation, (2) mixed and (3) drop out | Not reported | Mortality | Cox proportional hazards regression Models for NC categories were used to evaluate relationship between NC and mortality. Not clear whether potential confounding effects of diffuse cutaneous SSc and ILD adjusted for in regression modelling. | There were 331 deaths (number of deaths in subgroup with NC data not reported). No relationship between NC patterns and survival were identified (trend for higher mortality with increased capillary drop out). A significant Cox Proportional hazards regression ratio of 1.60 (95% CI 1.06-2.42; P=0.027) was observed when categorical NC patterns were treated as continuous variables (scores 0-3) but not clear whether diffuse cutaneous SSc and ILD incorporated in hazards model. |
| Smith et al. 2011 [36] | Belgium | P L | NC predictors for future DU at 6-12 months. | 71 | 6 subjects lost to follow-up | LeRoy and Medsger classification criteria for early SSc [7] | Nil | 2nd-5th digits of both hands. x200 magn. | Mean cap loss at 32 sites (4/finger) based on number of capillaries in 1mm (0, normal, 1, <33% loss [9-12/mm], 2, 33-66% loss [6-9/mm], 3, >66% loss [<6/mm]). | Yes | Digital trophic changes using DSS peripheral vascular domain | ROC curve analysis was undertaken to assess optimal cut-offs for mean score of capillary loss (calculated at 32 sites, 16 sites, 8 sites and at 4 sites) for predicting the presence of digital trophic lesions at baseline and at 6-12 months follow-up. | Digital trophic changes were present in 19/71 patients at baseline and 22/65 patients at 6-12 month follow up. The NC image analysis was blinded but the choice of images may have been influenced by the presence of digital ischaemic lesions at the time of image acquisition. The analysis for predicting digital trophic lesions at 6-12 months did not adjust for important confounders such as DU at baseline. Statistical analysis not optimal for stated aim of investigating predictive value of NC changes. |
| Smith et al. 2012 [37] | Belgium | P L | NC predictors of future organ complications of SSc at 18-24 months. | 66 | 8 subjects lost to follow-up | LeRoy and Medsger classification criteria for early SSc [7] | Nil | 2nd-5th digits of both hands at x200 magn. (x4 1mm fields per digit) | Global qualitative assessment of 32 NC images classified as "normal", "early", "active" and "late" changes [5]. | Yes | 9 organs systems using the DSS (0-4 semi-quantitative score) at 18-24 months. | Logistic regression analysis was undertaken, converting NVC assessment into a continuous variable for most analyses (unless rates of organ involvement were low). | Reported significant association between baseline NVC severity (1-4 as continuous scale) and presence of severe (category 2-4 of DSS) for peripheral vascular and pulmonary (ILD rather than PAH) organ systems on 18-24 month assessment. Baseline organ severity not reported and not clear whether organ manifestations differed at time of NVC assessment and assessment at 18-24 months. No adjustment for confounding effects of disease severity at baseline. Baseline NVC associated but not necessarily "predictive" of future organ involvement. |
| Sebastiani et al. 2012 [33] | Italy | P L | Multicentre study to validate CSURI in predicting new DU/DU healing at 3 months. | 242 | 13 excluded as lack of megacapillaries (a feature of advanced disease) | 1980 ARA classification criteria for SSc and at least 1 megacapillary to be able to calculate the CSURI | Use of bosentan. | 2nd-5th digits on both hands at x200 magn. | A single image from the 8 digits used to calculate CSURI score was calculated using the formula: $DxM/N^2$ (as described above Sebastiani 2009) | Not reported | New DU or persistence of old/recurrence DU at 3 months. | ROC curve analysis examining optimal cut-off and predictive accuracy for CSURI when comparing patients with new/non-healing DU. | 39 patients had DU at baseline (persisted at 3 months in 21 subjects) and 36 patients developed new DU at 3 months. A CSURI cut-off value of 2.96 (near identical to earlier study) had a specificity of 81.4% and sensitivity of 92.98% for predicting new DU or non-healing of pre-existing ulcers at 3 months. PPV of CSURI 62.3% with NPV of 97.2%. The analytical approach (ROC curve analyses) does not allow for contribution of potentially important confounders e.g. disease duration and history of DU. The presence of DU at baseline might have influenced NC image acquisition. |
| Kayser et al. 2013 [44] | Brazil | R L | Risk factors for mortality in SSc | 205 | 70 lost to follow-up (excluded from analysis) | 1980 ARA classification criteria for SSc (seen 2001 to 2009. | Overlap symptoms to other CTD e.g. SLE, RA. | All 10 digits. 10-25x magn. | (i) N of cap loops/mm (ii) avascular score (graded 0-3 according to number of avascular areas per digit) (iii) N of enlarged caps (>4x normal size) (iv) N of giant NC (>10x normal) | Yes | Mortality during mean 6.5 years of follow up. | Univariate Cox proportional hazards models used to examine relationship between variables and mortality. A multivariate model built using variables identified at univariate analysis was created. Survival analysis using Kaplan Meier Curve analysis was undertaken to analyse survival using an arbitrary avascular score cut-off of 1.5. | There were only 12 deaths (5 of which were attributable to SSc-lung involvement) over a variable follow-up ranging from 1-22 years. Three patients died secondary to PAH. Univariate analysis identified male gender, number of giant capillary loops and avascular score >1.5 as associated with increased mortality. After multivariate analysis, only an average digital avascular score of >1.5 was associated with a 2-fold increased mortality. Important potential confounders e.g. organ involvement such as PAH at presentation not accounted in analysis. |
| Sebastiani et al. 2013 [32] | Italy | P L | Multicentre study exploring predictive value of CSURI in | 176 | None reported | 1980 ARA classification criteria for SSc | Bosentan use | 2nd-5th digits on both hands at x200 magn. | A single image from the 8 digits was used to calculate CSURI score | Not reported | New DU or persistence of old/recurren | ROC curve analysis examining optimal cut-off and predictive accuracy for CSURI when comparing patients with new/non-healing DU | The study examined whether image width (1.33 – 1.70mm) influenced predictive value of the CSURI. The optimal CSURI cut-offs (using ROC curve analysis) to predict new DU occurrence or non-healing of existing DUs differed between cameras (2.13-3.81). The analytical approach did |

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | predicting new DU/DU healing. | | | | | (using 3 different NC devices). | using the formula: $DxM/N^2$ (as described above)) | | t lesions at 3 months. | | not allow for potentially important confounders e.g. prior history of DU in predicting new DU occurrence or non-healing of existing DU. The presence of DU at baseline might have influenced NC image acquisition |
| Smith et al. 2013 [38] | Belgium & Italy | P L | NC predictors of future organ complications at 18-24 months | 148 SSc patients (66 Belgium patients, 82 Italian patients). | Insufficient follow up data on 14 subjects | Clinician diagnosis of SSc. Classification criterial not specified. | Nil | $2^{nd}$-$5^{th}$ digits both hands at x200 magn. (x4 1mm images per digit). | Global qual assessment of 32 NC images classified as "nonspecific", "normal", "early", "active" and "late" changes [5]. | Yes | 9 organ systems of DSS (0-4 score) and/or new PAH/ILD. | Simple and multiple (adjusting for disease duration, disease subset and vasoactive medication) logistic regression analysis. | NC changes at baseline (early, active and late) were associated with a significantly increased risk of worsening of the Peripheral Vascular Score to >2 points compared with patients with normal/non-specific changes (OR 2.96 [95% CI 1.45-7.05], p=0.002 for early versus normal NC pattern) on multiple logistic regression analysis. Analysis did not adequately control for important potential confounders at baseline assessment. |
| Violliot et al. 2015 [39] | Belgium and Italy | P L | Predictors of (e.g. exercise TTE, serum BNP and NC) associated with "CV events" during mean follow-up of 27 ± 18 months | 65 | 15 excluded (unquantifiable SPAP on TTE, mitral regurgitation and/or CHD). | Clinician diagnosis of SSc. Classification criterial not specified. | Unquantifiable sPAP, IHD, valvular disease, unable to perform exercise TTE | $2^{nd}$-$5^{th}$ digits both hands at x200 magn. | Global qualitative assessment of images classified as "normal", "early", "active" and "late" changes [5]. | Not reported | A heterogeneous range of "CV-events" | Relationship between "CV events" (comprising death/hospitalization due to left or right heart failure, atrial or ventricular arrhythmia, resting PAH confirmed on RHC, cerebrovascular attack or peripheral ulcer requiring hospitalisation) and NC changes (plus serum BNP, demographics, resting and exercise TTE data) evaluated using simple linear regression. Independent predictors of CV events assessed using multiple logistic regression analyses. | Heterogeneous mix of CV events in 9/50 patients. Patients with "CV events" were older, higher baseline BNP levels and differences in baseline TTE parameters e.g. higher baseline sPAP, higher exercise sPAP and larger LA area. Patients with future CV events were more likely to have late NC changes and less likely to have early NC changes. The relationship between an NC "active"/"late" changes and future CV events persisted after adjustment for age and sex. Multiple logistic regression analysis adjusted for age and gender but did not adjust for other important potential confounders (e.g. BNP, exercise TTE parameters etc.) which had an incremental effect on future CV events. There was limited description of baseline features, recruitment period, length of follow-up, blinding etc. |
| Silva et al. 2015 [34] | Portugal | P L | Risk factors (including NC) for further DU recurrence in patients with history of DU (group 1) or $1^{st}$ DU (group 2) over 3 years | 77 (group 1 n=38 and group 2 n=39) | Nil | 2013 ACR/EULAR classification criteria | Smokers, diabetes, hyperlipidaemia, history of MI and bosentan use | 200x magn. Number of digits assessed and images obtained/assessed not supplied. | Global qualitative assessment of images classified as "normal", "early", "active" and "late" changes [5]. | Not reported | DU reoccurrence (group 1) or $1^{st}$ DU event (group 2) in the 3 year follow up period | Univariate and multivariate cox regression analyses used to evaluate association between baseline variables (including NC pattern) and DU reoccurrence (Group 1) or $1^{st}$ DU event (Group 2). | 30/38 (79%) of patients in group 1 had DU reoccurrence within 3 years. Univariate and multivariate analyses both identified NVC "late" pattern as a predictor of DU reoccurrence in group 1 patients (HR 2.49 [95%CI 1.03-6.04] on multivariate cox regression analysis). 10/39 patients in group 2 experienced their first DU in 3-year clinical follow up. Univariate and multivariate analyses both identified NVC "late" pattern as a predictor of $1^{st}$ DU occurrence in group 2 patients (HR 13.38, [95%CI 2.10-85.38] on multivariate cox regression analysis). Limited description of NC methods and image analysis. Inadequate adjustment for confounders. |
| Manfredi et al. 2015 [30] | Italy | P L | Multicentre study examining risk factors for new DU over 6 months | 219 | Nil | 2013 ACR/EULAR classification criteria | Use of Bosentan | $2^{nd}$-$5^{th}$ digits of both hands. 200x magn. 1 image/finger obtained. | A single image from the 8 digits used to calculate CSURI using the formula: $DxM/N^2$ (as described above Sebastiani 2009) | Not reported | The occurrence of new DU over 6-month follow-up | Univariate and multivariate analysis to examine predictive value of baseline NVC changes at predicting new DU in 6 months of follow up. A prediction chart was devised using parameters identified in multivariate logistic regression analysis. | Multivariate analysis identified the CSURI cut off of 2.96 as highly predictive for new DU at 6 months. Approximately 84 pts had recent history/active DU at baseline and a similar number had a CSURI of >2.96 suggesting possible overlap. The high odds ratio (99.95 (95%CI 20.03-498.86, P<0.001) at multivariate analysis suggests quasi-complete separation owing to strong relationship between CSURI and baseline DU. |
| Violliot et al. 2016 [40] | Belgium and Italy | P L | Risk factors (resting TTE parameters, BNP, NC, PFTs) for PAH or EIPH 25 ±15 months follow-up | 68 recruited (2008-2012). 40 included in analysis. | 5 refused to take part. 2 lost to follow-up. 21 excluded. | Clinician diagnosis of SSc | CHD, valve disease, PAH, no exercise test or no TR on TTE | $2^{nd}$-$5^{th}$ digits both hands at x200 magn. | Global qualitative assessment of images classified as 0, "normal"; 1, "early"; 2, "active" and 3, "late" changes [5]. | Not reported | PAH defined as estimated TTE sPAP >35mmHg. EIPH as estimated TTE sPAP >50mmHg | Univariate and multivariable logistic regression and cox proportional hazards models were performed to define predictors of new onset PAH. | EIPH occurred in 11 patients (28%) during follow up. Patients developing EIPH had higher levels of baseline BNP, sPAP, mPAP and sPAP on TTE. After adjusting for age, patients with NVC>2 ("active"/"late") were more likely to develop PAH during follow up (hazard ratio 9.1, 95% CI 1.1-74.8; P=0.04). Multivariate analysis of more than 2 variables was not possible (small study size and low event rate) raising concerns regarding type II error. The exclusion criteria and reliance on TTE parameters may also reduce the generalisability of findings across unselected SSc cohorts. |
| Baron et al. 2016 [41] | Canada and Mexico | R L | Multicentre longitudinal study examining risk factors for calcinosis | 1305 patients recruited at 15 sites (2004 to 2015) | 500 pts excluded from analysis | Clinician diagnosis of SSc antibody data and calcinosis assessment at baseline | Baseline calcinosis (n=300) or no follow-up (n=200) | Dermatoscope used for NC. Details not reported. | Physician recorded the presence of capillary drop-out areas. | Not reported | Development of calcinosis during mean follow up of 3.8years | Univariate and multivariate analysis examined predictors of calcinosis. A cox proportional hazards model examined the association between baseline markers of tissue ischaemia (including capillary drop out) and future development of calcinosis. | Capillary dropout at baseline was considered a significant univariate predictor of calcinosis during follow-up (p=0.063). In a sensitivity analysis, capillary drop out at baseline was also associated with the development of calcinosis (HR 1.46, 95% CI: 1.08, 1.99). Limited details were provided regarding the NC image acquisition and analysis. |
| Cutolo et al. 2016 [29] | 59 SSc centres across 14 countries (12 in Europe, Turkey and Israel) | P L | Multicentre study examining risk factors (particularly NC) for new DU occurrence over 6 months | 623 patients: 468 in group 1 and 123 in group 2) | No evaluable DU outcome data for 32 patients. | Classification criteria for early SSc [7] 1. history of DU or DU at enrolment (group 1) 2. Disease duration of <2 years (group 2) | Inability to do NC, SSc sine scleroderma, previous HSCT or clinical trial within 3 months | $2^{nd}$-$5^{th}$ digits both hands at x200 magn. capturing 2 adjacent fields extending over 1mm at the midline. | Qualitative assessment classified as "normal", "early", "active" and "late" changes [5]. Quant assessment of n of caps, giant caps, irregularly enlarged caps, micro-haemorrhages and neoangiogenesis. Maximal cap diameter was measured. | No | Occurrence of new DU reported during monthly monitoring phone calls to patients (confirmed at study visit). | Stratified analysis undertaken focussing on patients with history of DU (n=468). Logistic regression modelling (including linear and quadratic functional relationships) examined discriminatory ability of co-variables for the occurrence of new DU. Strength of association reported using odds ratio and discriminatory performance assessed using ROC AUC analyses. Multiple logistic regression modelling used to identify best-performing clinical and NC risk factors for DU occurrence. | There were 103 new DU in group 1 patients and only 5 new DU in the group 2 patients. The presence and number of DU present at baseline was the strongest predictor of new DU during follow up. Qualitative NC pattern was associated with new DU occurrence when comparing "late" versus "normal/early". The mean number of capillaries was significant lower in patients who experienced new DU compared with those who did not. The final multiple logistic regression identified 3 risk factors for DU occurrence that included: 1) Mean number of capillaries/mm in middle finger of dominant hand, 2) the number of DUs at enrolment visit and 3) the presence of critical digital ischaemia at enrolment. The stratified analysis reduces the generalisability of the findings in patients without a history of DU. |

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Avouac et al. 2017 [28] | France, Italy, Belgium | P L | Study examining evolution of NC changes and relationship with disease progression over 3 year follow-up. | 146 Enrolled (2011 to 2012). | 6 patients died during follow up. 15 patients lost to follow-up. | 2013 ACR/EULAR classification criteria | Nil | $2^{nd}$-$5^{th}$ digits both hands at x200 magn. Two consecutive fields extending over 1mm at midline. NC repeated annually for 3 years. | Quantitative assessment of the mean number of a) capillaries per mm, b) giant capillaries per mm, c) microhaemorrhages per mm and d) neoangiogenesis per mm. A qualitative grading of NVC changes as “early”, “active” and “late” NVC patterns.[5] | Yes | “Disease progression” during 3 year follow-up. | “Disease progression” defined as progression of modified RSS (≥30% and ≥5 point increase), occurrence of DU, new onset PAH on RHC, cardiac involvement (LVEF to <50% on TTE), renal crisis, progression of ILD (new onset ILD on HRCT or >10% fall in FVC) or worsening of the DSS (from category 1-2 to 3-4 for any organ system). An analysis of the predictive value of baseline NC changes using univariate Cox proportional hazards models followed by multivariate analysis including all variables with P<0.1 from univariate analysis. | Baseline capillary loss was predictor of overall disease progression, occurrence of new DU, lung vascular progression, skin progression and worsening of the DSS at univariate analysis but not multivariate analysis. In contrast, neoangiogenesis at baseline retained a significant association with the lung vascular progression (HR 11.12, 95%CI 1.19-103.79, p=0.036) on multivariate analysis adjusting for confounders. Evolution of NC abnormalities was more important; progression in the number of giant caps was protective for occurrence of new DU and overall disease progression. Progression of cap loss was independently associated with overall disease progression, occurrence of new DU, lung vascular progression, skin progression and worsening of the DSS. Progression to “late” NVC changes independently associated with new DU, development of PAH, progression of skin fibrosis and worsening of the DSS. |

ACR, American College of Rheumatology; AF, atrial fibrillation; AUC, area under curve; ARA, Australian Rheumatology Association; BNP, Brain Natriuretic Peptide; Cap, capillary; CHD, Coronary Heart Disease; CI, Confidence Intervals; CSURI, capillaroscopic skin ulcers risk index; CV, Cardiovascular; DU, Digital Ulcers; DSS, Disease Severity Score; EULAR, European League Against Rheumatism; EIPH, Exercise-induced Pulmonary Hypertension; FVC, Forced Vital Capacity; IHD, Ischaemic Heart Disease; ILD, Interstitial Lung Disease; Hb, Haemoglobin; HDL, High-density Lipoprotein; HRCT, High Resolution Computed Tomography; HR, Hazards Ratio; HSCT, Haematopoietic Stem Cell Transplantation; LA, Left Atrium; LVEF, Left Ventricular Ejection Fraction; MI, Myocardial Infarction; mPAP, mean Pulmonary Arterial Pressure; NC, Nailfold Capillaroscopy; NPV, Negative Predictive Value; OR, Odds Ratio; P, Prospective; Pts, patients; PAH, Pulmonary Arterial Hypertension; PF, Pulmonary Fibrosis; PPV, Positive Predictive Value; Qual, qualitative; Quant, quantitative; R, Retrospective; RA, Rheumatoid Arthritis; RHC, Right Heart Catheterisation; ROC, Receiver Operating Characteristic; RSS, Rodnan Skin Score; SLE, Systemic Lupus Erythematosus; sPAP, systolic Pulmonary Arterial Pressure; SSc, Systemic Sclerosis; TIA, Transient Ischaemic Attack; TTE, Transthoracic Echocardiogram; vTR, velocity Tricuspid Regurgitation.

**Table 2. Risk of bias assessment in studies reporting the prognostic value of nailfold capillaroscopy in systemic sclerosis using the QUality in Prognosis Studies (QUIPS) tool.**

| Author & Year of Publication | Study Participation | Study Attrition | Prognostic Factor Measurement | Outcome Measurement | Study Confounding | Statistical Analysis & Reporting |
|---|---|---|---|---|---|---|
| Scussel-Lonzetti et al. 2002 [42] | Low | Low | Low | Low | High | High |
| Simeón et al. 2003 [35] | Low | Low | Moderate | Low | Moderate | Moderate |
| Alivernini et al. 2009 [27] | Low | Low | High | Moderate | Low | High |
| Sebastiani et al. 2009 [31] | Moderate | Low | Moderate | Moderate | High | High |
| Hissaria et al. 2011 [43] | Low | Low | Low | Low | High | Moderate |
| Smith et al. 2011 [36] | Low | Low | Low | Low | High | High |
| Smith et al. 2012 [37] | Low | Low | Low | Low | High | Moderate |
| Sebastiani et al. 2012 [33] | Low | Low | Moderate | Moderate | High | High |
| Kayser et al. 2013 [44] | Low | High | Low | Low | High | Low |
| Sebastiani et al. 2013 [32] | Low | Low | Moderate | Moderate | High | High |
| Smith et al. 2013 [38] | Moderate | Moderate | Low | Low | Low | Moderate |
| Violliot et al. 2015 [39] | Moderate | Low | Moderate | Low | High | Moderate |
| Silva et al. 2015 [34] | Low | Low | Moderate | Low | Moderate | Low |
| Manfredi et al. 2015[30] | Moderate | Low | Moderate | Low | High | Moderate |
| Violliot et al. 2016 [40] | Moderate | Moderate | Moderate | Low | High | Moderate |
| Baron et al. 2016 [41] | Low | Low | Moderate | Low | Low | Low |
| Cutolo et al. 2016 [29] | Moderate | Low | Low | Moderate | Low | Low |
| Avouac et al. 2017 [28] | Low | Low | Low | Low | Low | Low |
| **Totals:**<br>Low (58)<br>Mod (31)<br>High (19) | <br>12<br>6<br>0 | <br>15<br>2<br>1 | <br>8<br>9<br>1 | <br>13<br>5<br>0 | <br>5<br>2<br>11 | <br>5<br>7<br>6 |

**The prognostic value of nailfold capillaroscopic abnormalities in systemic sclerosis: A systematic literature review**

**Dolcie Paxton BSc & John D Pauling BMedSci BMBS PhD FRCP**

**Supplementary Material 1. Details of full search criteria applied in MEDLINE and EMBASE databases for study identification.**

**MEDLINE search**

(Nailfold videocapillaroscopy or nailfold capillary microscopy or capillaroscopy or microscopic angioscopy) AND (scleroderma, systemic or scleroderma or systemic scleroderma or systemic sclerosis or CREST) AND (cohort studies or Longitudinal studies or case-control studies or follow-up studies or retrospective studies or cross-sectional studies or prospective studies or incidence or mortality or follow-up studies or prognos*or predict* or course or prognostic or prognosis or progression or future or development or treatment outcome or disease-free survival or treatment failure or morbidity or prevalence or survival rate or cause of death or survival analysis)

The expanded search is detailed below:

((nailfold[All Fields] AND ("microscopic angioscopy"[MeSH Terms] OR ("microscopic"[All Fields] AND "angioscopy"[All Fields]) OR "microscopic angioscopy"[All Fields] OR "videocapillaroscopy"[All Fields])) OR (nailfold[All Fields] AND ("capillaries"[MeSH Terms] OR "capillaries"[All Fields] OR "capillary"[All Fields]) AND ("microscopy"[MeSH Terms] OR "microscopy"[All Fields])) OR ("microscopic angioscopy"[MeSH Terms] OR ("microscopic"[All Fields] AND "angioscopy"[All Fields]) OR "microscopic angioscopy"[All Fields] OR "capillaroscopy"[All Fields]) OR ("microscopic angioscopy"[MeSH Terms] OR ("microscopic"[All Fields] AND "angioscopy"[All Fields]) OR "microscopic angioscopy"[All Fields])) AND (("scleroderma, systemic"[MeSH Terms] OR ("scleroderma"[All Fields] AND "systemic"[All Fields]) OR "systemic scleroderma"[All Fields] OR ("scleroderma"[All Fields] AND "systemic"[All Fields]) OR "scleroderma, systemic"[All Fields]) OR ("scleroderma, systemic"[MeSH Terms] OR ("scleroderma"[All Fields] AND "systemic"[All Fields]) OR "systemic scleroderma"[All Fields] OR "scleroderma"[All Fields] OR "scleroderma, localized"[MeSH Terms] OR ("scleroderma"[All Fields] AND "localized"[All Fields]) OR "localized scleroderma"[All Fields]) OR ("scleroderma, systemic"[MeSH Terms] OR ("scleroderma"[All Fields] AND "systemic"[All Fields]) OR "systemic scleroderma"[All Fields] OR ("systemic"[All Fields] AND "scleroderma"[All Fields])) OR ("scleroderma, systemic"[MeSH Terms] OR ("scleroderma"[All Fields] AND "systemic"[All Fields]) OR "systemic scleroderma"[All Fields] OR ("systemic"[All Fields] AND "sclerosis"[All Fields]) OR "systemic sclerosis"[All Fields]) OR ("crest syndrome"[MeSH Terms] OR ("crest"[All Fields] AND "syndrome"[All Fields]) OR "crest syndrome"[All Fields] OR "crest"[All Fields])) AND (("cohort studies"[MeSH Terms] OR ("cohort"[All Fields] AND "studies"[All Fields]) OR "cohort studies"[All Fields]) OR ("longitudinal studies"[MeSH Terms] OR ("longitudinal"[All

Fields] AND "studies"[All Fields]) OR "longitudinal studies"[All Fields]) OR ("case-control studies"[MeSH Terms] OR ("case-control"[All Fields] AND "studies"[All Fields]) OR "case-control studies"[All Fields] OR ("case"[All Fields] AND "control"[All Fields] AND "studies"[All Fields]) OR "case control studies"[All Fields]) OR ("follow-up studies"[MeSH Terms] OR ("follow-up"[All Fields] AND "studies"[All Fields]) OR "follow-up studies"[All Fields] OR ("follow"[All Fields] AND "up"[All Fields] AND "studies"[All Fields]) OR "follow up studies"[All Fields]) OR ("retrospective studies"[MeSH Terms] OR ("retrospective"[All Fields] AND "studies"[All Fields]) OR "retrospective studies"[All Fields]) OR ("cross-sectional studies"[MeSH Terms] OR ("cross-sectional"[All Fields] AND "studies"[All Fields]) OR "cross-sectional studies"[All Fields] OR ("cross"[All Fields] AND "sectional"[All Fields] AND "studies"[All Fields]) OR "cross sectional studies"[All Fields]) OR ("prospective studies"[MeSH Terms] OR ("prospective"[All Fields] AND "studies"[All Fields]) OR "prospective studies"[All Fields]) OR ("epidemiology"[Subheading] OR "epidemiology"[All Fields] OR "incidence"[All Fields] OR "incidence"[MeSH Terms]) OR ("mortality"[Subheading] OR "mortality"[All Fields] OR "mortality"[MeSH Terms]) OR ("follow-up studies"[MeSH Terms] OR ("follow-up"[All Fields] AND "studies"[All Fields]) OR "follow-up studies"[All Fields] OR ("follow"[All Fields] AND "up"[All Fields] AND "studies"[All Fields]) OR "follow up studies"[All Fields]) OR ((prognose[All Fields] AND *or[All Fields]) AND (predict[All Fields] OR predict'[All Fields] OR predict''[All Fields] OR predict's[All Fields] OR predict7[All Fields] OR predicta[All Fields] OR predictab[All Fields] OR predictabe[All Fields] OR predictabel[All Fields] OR predictabilities[All Fields] OR predictability[All Fields] OR predictability'[All Fields] OR predictability''[All Fields] OR predictabilty[All Fields] OR predictable[All Fields] OR predictable'[All Fields] OR predictables[All Fields] OR predictablility[All Fields] OR predictablity[All Fields] OR predictably[All Fields] OR predictabuity[All Fields] OR predictad[All Fields] OR predictal[All Fields] OR predictalbe[All Fields] OR predictand[All Fields] OR predictands[All Fields] OR predictaquatic[All Fields] OR predictar[All Fields] OR predictate[All Fields] OR predictated[All Fields] OR predictative[All Fields] OR predictbias[All Fields] OR predictcancer[All Fields] OR predictcing[All Fields] OR predictd[All Fields] OR predictdr[All Fields] OR predicte[All Fields] OR predicted[All Fields] OR predicted'[All Fields] OR predictedfev[All Fields] OR predictedfrom[All Fields] OR predictedindices[All Fields] OR predictedinteractions[All Fields] OR predictedl[All Fields] OR predictedly[All Fields] OR predictedmore[All Fields] OR predictedness[All Fields] OR predictedproperties[All Fields] OR predicteds[All Fields] OR predictedx100[All Fields] OR predictee[All Fields] OR predictees[All Fields] OR predictek[All Fields] OR predicter[All Fields] OR predicters[All Fields] OR predictet[All Fields] OR predicteur[All Fields] OR predicteurs[All Fields] OR predictfold[All Fields] OR predictfurors[All Fields] OR predicthaplo[All Fields] OR predicther[All Fields] OR predicthospital[All Fields] OR predicti[All Fields] OR predictia[All Fields] OR predictiable[All Fields] OR predictibilidad[All Fields] OR predictibilitate[All Fields] OR predictibilite[All Fields] OR predictibility[All Fields] OR predictible[All Fields] OR predictically[All Fields] OR predictice[All Fields] OR predictictive[All Fields] OR predicticvity[All Fields] OR predictie[All Fields] OR predictied[All Fields] OR predictief[All Fields] OR predictieformules[All Fields] OR predictiemodel[All Fields] OR predictiemodellen[All Fields] OR predictieregels[All Fields] OR predicties[All Fields] OR predictieve[All Fields] OR predictif[All Fields] OR predictif's[All Fields] OR predictifs[All Fields] OR predictim[All Fields] OR predictin[All Fields] OR predictinf[All Fields] OR predicting[All Fields] OR predicting'[All Fields] OR predicting14[All Fields] OR predictingdti[All Fields] OR predictinginteractions[All Fields] OR

predictingprognosis[All Fields] OR predictingthe[All Fields] OR predictintegral[All Fields] OR predictio[All Fields] OR predictiom[All Fields] OR prediction[All Fields] OR prediction'[All Fields] OR prediction's[All Fields] OR prediction36[All Fields] OR predictionable[All Fields] OR predictional[All Fields] OR predictionalgorithms[All Fields] OR predictioncenter[All Fields] OR predictioncomputations[All Fields] OR predictionequations[All Fields] OR predictiong[All Fields] OR predictioning[All Fields] OR predictionnel[All Fields] OR predictionof[All Fields] OR predictions[All Fields] OR predictions'[All Fields] OR predictionst[All Fields] OR predictionswere[All Fields] OR predictior[All Fields] OR predictis[All Fields] OR predictit[All Fields] OR predictition[All Fields] OR predictitive[All Fields] OR predictits[All Fields] OR predictiv[All Fields] OR predictiva[All Fields] OR predictivas[All Fields] OR predictive[All Fields] OR predictive'[All Fields] OR predictiveanalytics[All Fields] OR predictiveathletes[All Fields] OR predictiveathletes'[All Fields] OR predictivebio[All Fields] OR predictivefactor[All Fields] OR predictivefactors[All Fields] OR predictivefactorsfor[All Fields] OR predictivehealth[All Fields] OR predictivein[All Fields] OR predictively[All Fields] OR predictively'[All Fields] OR predictivemodel[All Fields] OR predictiveness[All Fields] OR predictiveness'[All Fields] OR predictivenets[All Fields] OR predictivenetworks[All Fields] OR predictiveof[All Fields] OR predictives[All Fields] OR predictivesolutions[All Fields] OR predictivetrade[All Fields] OR predictivevalue[All Fields] OR predictivevalues[All Fields] OR predictivi[All Fields] OR predictivie[All Fields] OR predictivism[All Fields] OR predictivist[All Fields] OR predictivite[All Fields] OR predictivites[All Fields] OR predictivities[All Fields] OR predictivity[All Fields] OR predictivo[All Fields] OR predictivos[All Fields] OR predictivre[All Fields] OR predictivy[All Fields] OR predictly[All Fields] OR predictlymph[All Fields] OR predictmalaga[All Fields] OR predictmdxtrade[All Fields] OR predictment[All Fields] OR predictnls[All Fields] OR predictol[All Fields] OR predictome[All Fields] OR predictomics[All Fields] OR predicton[All Fields] OR predictons[All Fields] OR predictor[All Fields] OR predictor'[All Fields] OR predictor''[All Fields] OR predictor's[All Fields] OR predictor2[All Fields] OR predictora[All Fields] OR predictoras[All Fields] OR predictore[All Fields] OR predictoren[All Fields] OR predictores[All Fields] OR predictoresde[All Fields] OR predictorfor[All Fields] OR predictorganic[All Fields] OR predictormultiple[All Fields] OR predictorof[All Fields] OR predictorr[All Fields] OR predictors[All Fields] OR predictors'[All Fields] OR predictors'of[All Fields] OR predictors,[All Fields] OR predictors1[All Fields] OR predictorsand[All Fields] OR predictorsof[All Fields] OR predictorsqbetween[All Fields] OR predictorstate[All Fields] OR predictortrade[All Fields] OR predictorvsl2[All Fields] OR predictorxgroup[All Fields] OR predictory[All Fields] OR predictove[All Fields] OR predictovos[All Fields] OR predictox[All Fields] OR predictpatientevents[All Fields] OR predictprotei[All Fields] OR predictprotein[All Fields] OR predictprotein'[All Fields] OR predictree[All Fields] OR predictregulon[All Fields] OR predictreturn[All Fields] OR predictrive[All Fields] OR predicts[All Fields] OR predictsa[All Fields] OR predictsgrna[All Fields] OR predictsnon[All Fields] OR predictsnp[All Fields] OR predictsnp1[All Fields] OR predictsnp2[All Fields] OR predictt[All Fields] OR predicttga[All Fields] OR predictthe[All Fields] OR predicttion[All Fields] OR predicttive[All Fields] OR predicttoxicity[All Fields] OR predicttrade[All Fields] OR predictve[All Fields] OR predictyate[All Fields])) OR course[All Fields] OR prognostic[All Fields] OR ("prognosis"[MeSH Terms] OR "prognosis"[All Fields]) OR ("disease progression"[MeSH Terms] OR ("disease"[All Fields] AND "progression"[All Fields]) OR "disease progression"[All Fields] OR "progression"[All Fields]) OR ("forecasting"[MeSH Terms] OR "forecasting"[All Fields] OR "future"[All Fields]) OR ("growth

and development"[Subheading] OR ("growth"[All Fields] AND "development"[All Fields]) OR "growth and development"[All Fields] OR "development"[All Fields]) OR ("treatment outcome"[MeSH Terms] OR ("treatment"[All Fields] AND "outcome"[All Fields]) OR "treatment outcome"[All Fields]) OR ("disease-free survival"[MeSH Terms] OR ("disease-free"[All Fields] AND "survival"[All Fields]) OR "disease-free survival"[All Fields] OR ("disease"[All Fields] AND "free"[All Fields] AND "survival"[All Fields]) OR "disease free survival"[All Fields]) OR ("treatment failure"[MeSH Terms] OR ("treatment"[All Fields] AND "failure"[All Fields]) OR "treatment failure"[All Fields]) OR ("epidemiology"[Subheading] OR "epidemiology"[All Fields] OR "morbidity"[All Fields] OR "morbidity"[MeSH Terms]) OR ("epidemiology"[Subheading] OR "epidemiology"[All Fields] OR "prevalence"[All Fields] OR "prevalence"[MeSH Terms]) OR ("survival rate"[MeSH Terms] OR ("survival"[All Fields] AND "rate"[All Fields]) OR "survival rate"[All Fields]) OR ("cause of death"[MeSH Terms] OR ("cause"[All Fields] AND "death"[All Fields]) OR "cause of death"[All Fields]) OR ("survival analysis"[MeSH Terms] OR ("survival"[All Fields] AND "analysis"[All Fields]) OR "survival analysis"[All Fields]))

**EMBASE search**

Nailfold videocapillaroscopy or nailfold capillary microscopy or capillaroscopy or microscopic angioscopy

AND

scleroderma, systemic or scleroderma or systemic scleroderma or systemic sclerosis or CREST

AND

cohort studies or Longitudinal studies or case-control studies or follow-up studies or retrospective studies or cross-sectional studies or prospective studies or incidence or mortality or follow-up studies or prognos*or predict* or course or prognostic or prognosis or progression or future or development or treatment outcome or disease-free survival or treatment failure or morbidity or prevalence or survival rate or cause of death or survival analysis

## Supplementary Materials 2. The "QUality In Prognosis Studies" (QUIPS) risk-of-bias tool.

The following criteria were used to rate the risk of bias across the 6 domains of the QUIPS tool. Adapted from [25]. PF, prognostic factor.

| Bias Domain | Optimal Study Description | Prompting Items and Considerations | Final rating |
|---|---|---|---|
| **Study Participation** | The study sample adequately represents the population of interest | a. Adequate participation in the study by eligible persons<br>b. Description of the source population or population of interest<br>c. Description of the baseline study sample<br>d. Adequate description of the sampling frame and recruitment<br>e. Adequate description of the period and place of recruitment<br>f. Adequate description of inclusion and exclusion criteria | **High bias**: The relationship between the PF and outcome is very likely to be different for participants and eligible nonparticipants |
| | | | **Moderate bias:** The relationship between the PF and outcome may be different for participants and eligible nonparticipants |
| | | | **Low bias:** The relationship between the PF and outcome is unlikely to be different for participants and eligible nonparticipants |
| **Study Attrition** | The study data available (i.e., participants not lost to follow-up) adequately represent the study sample | a. Adequate response rate for study participants<br>b. Description of attempts to collect information on participants who dropped out<br>c. Reasons for loss to follow-up are provided<br>d. Adequate description of participants lost to follow-up<br>e. There are no important differences between participants who completed the study and those who did not | **High bias:** The relationship between the PF and outcome is very likely to be different for completing and non-completing participants |
| | | | **Moderate bias:** The relationship between the PF and outcome may be different for completing and non-completing participants |
| | | | **Low bias:** The relationship between the PF and outcome is unlikely to be different for completing and non-completing participants |
| **Prognostic Factor (PF) Measurement** | The prognostic factor is measured in a similar way for all participants | a. A clear definition or description of the PF is provided<br>b. Method of PF measurement is adequately valid and reliable<br>c. Continuous variables are reported or appropriate cut points are used<br>d. The method and setting of measurement of PF is the same for all study participants<br>e. Adequate proportion of the study sample has complete data for the PF<br>f. Appropriate methods of imputation are used for missing PF data | **High bias:** The measurement of the PF is very likely to be different for different levels of the outcome of interest |
| | | | **Moderate bias:** The measurement of the PF may be different for different levels of the outcome of interest |
| | | | **Low bias:** The measurement of the PF is unlikely to be different for different levels of the outcome of interest |
| **Outcome Measurement** | The outcome of interest is measured in a similar way for all participants | a. A clear definition of the outcome is provided<br>b. Method of outcome measurement used is adequately valid and reliable<br>c. The method and setting of outcome measurement is the same for all study participants | **High bias:** The measurement of the outcome is very likely to be different related to the baseline level of the PF |
| | | | **Moderate bias:** The measurement of the outcome may be different related to the baseline level of the PF |
| | | | **Low bias:** The measurement of the outcome is unlikely to be different related to the baseline level of the PF |
| **Study Confounding** | Important potential confounding factors are appropriately accounted for | a. All important confounders are measured<br>b. Clear definitions of the important confounders measured are provided<br>c. Measurement of all important confounders is adequately valid and reliable<br>d. The method and setting of confounding measurement are the same for all study participants<br>e. Appropriate methods are used if imputation is used for missing confounder data<br>f. Important potential confounders are accounted for in the study design<br>g. Important potential confounders are accounted for in the analysis | **High bias:** The observed effect of the PF on the outcome is very likely to be distorted by another factor related to PF and outcome |
| | | | **Moderate bias:** The observed effect of the PF on outcome may be distorted by another factor related to PF and outcome |
| | | | **Low bias:** The observed effect of the PF on outcome is unlikely to be distorted by another factor related to PF and outcome |
| **Statistical Analysis and Reporting** | The statistical analysis is appropriate, and all primary outcomes are reported | a. Sufficient presentation of data to assess the adequacy of the analytic strategy<br>b. Strategy for model building is appropriate and is based on a conceptual framework or model<br>c. The selected statistical model is adequate for the design of the study<br>d. There is no selective reporting of results | **High bias:** The reported results are very likely to be spurious or biased related to analysis or reporting |
| | | | **Moderate bias:** The reported results may be spurious or biased related to analysis or reporting |
| | | | **Low bias:** The reported results are unlikely to be spurious or biased related to analysis or reporting |

**The prognostic value of nailfold capillaroscopic abnormalities in systemic sclerosis: A systematic literature review**

**Dolcie Paxton BSc & John D Pauling BMedSci BMBS PhD FRCP**

**Supplementary Materials 3. Details of inter-reviewer agreement during initial title and abstract review for eligible articles.**

Details of agreement between reviewers for the title and abstract review

| | **DP include** | **DP Exclude** | **Total** |
|---|---|---|---|
| **JP Include** | 18 | 7 | **25** |
| **JP Exclude** | 8 | 686 | **694** |
| **Total** | **26** | **693** | **719** |

**Number of observed agreements: 704 (97.91%)**

**Kappa = 0.6951 (95% CI 0.549-0.841)**
**SE of kappa=0.075**

Of the 15 citations for which there was disagreement:

5 papers were included in full text analysis (2 which JP had initially considered including)
10 papers were excluded (5 which JP had initially considered including)

This resulted in a total of 23 papers proceeding to full text review.

**The prognostic value of nailfold capillaroscopic abnormalities in systemic sclerosis: A systematic literature review**

**Dolcie Paxton BSc & John D Pauling BMedSci BMBS PhD FRCP**

Supplementary Material 4. Details of inter-reviewer agreement during QUIPS risk-of-bias assessment of articles taken forward to full data extraction

Details of agreement between reviewers for the QUIPS assessment

| | | **DP assessment** | | | **Totals** |
|---|---|---|---|---|---|
| **JP asssessment** | | Low | Mod | High | |
| | Low | 44 | 8 | 1 | 53 |
| | Mod | 17 | 18 | 4 | 39 |
| | High | 3 | 1 | 11 | 15 |
| **Totals** | | 64 | 27 | 16 | 107* |

* DP was undecided on 1 assessment and did not disclose an opinion during independent QUIPS assessment.

Number of observed agreements: 73 ( 68.22% of the observations)
Number of agreements expected by chance: 43.8 ( 40.92% of the observations)

Kappa= 0.462
SE of kappa = 0.075
95% confidence interval: From 0.314 to 0.610
The strength of agreement is considered to be 'moderate'.

The calculations above only consider exact matches between observers. As the categories (Low, Mod and High) are ordered, we also considered close matches. In other words, if one observer classifies a bias assessment as Moderate and the other rater High, this is closer than if one classifies one as Low and the other rater as High. The calculation of weighted kappa, below, assumes the categories are ordered and accounts for how far apart the two raters are. This calculation uses linear weights.

Weighted Kappa= 0.526

Assessed this way, the strength of agreement is considered to be 'moderate'.

A meeting was held between JP and DP on the 14th September to discuss dispcrepancies in agreement. The following decisions were agreed:

- For the manuscript that DP could not apply a grading, consensus was reached to apply JP grading
- For 5 gradings, it was agreed that we would apply the initial grading provided by DP

- For 28 gradings, it was agreed that we would apply the initial grading provided by JP
- For 1 grading, where there had been disagreement, we applied an entirely new grading (High when we had initially graded Moderate and Low respectively)
- For 3 gradings, where there had initially been agreement (both reviewers rated moderate), we applied a new grading (increased to high) based on further review of the study alongside similar assessments made for other studies.