Rheumatology, 2024, **63**, 1502–1506
https://doi.org/10.1093/rheumatology/kead337
Advance access publication 8 July 2023
Concise Report

British Society for Rheumatology RHEUMATOLOGY OXFORD

Clinical science

Assessment of digital perfusion as a surrogate outcome in Raynaud's phenomenon clinical trials

Alicia Guigui[1,2], Léa Liaigre[2], Marc Manceau[2], Olivier Gaget[2], Jean-Luc Cracowski[1,3], Sophie Blaise[1,4], Charles Khouri [1,2,3], Matthieu Roustit [1,2,*]

[1]Univ. Grenoble Alpes, Inserm, CHU Grenoble Alpes, HP2, Grenoble, France
[2]Univ. Grenoble Alpes, Inserm CIC1406, CHU Grenoble Alpes, Grenoble, France
[3]Pharmacovigilance Unit, Univ. Grenoble Alpes, CHU Grenoble Alpes, Grenoble, France
[4]Department of Vascular Medicine, Univ. Grenoble Alpes, CHU Grenoble Alpes, Grenoble, France

*Correspondence to: Matthieu Roustit, Unité de Pharmacologie Clinique, Centre d'Investigation Clinique de Grenoble—Inserm CIC1406, CHU Grenoble Alpes, CS 10217, 38043 Grenoble Cedex 09, France. E-mail: MRoustit@chu-grenoble.fr

Abstract

Objectives: Measurement of digital perfusion, sometimes coupled with a cold challenge, has been widely used as an objective outcome in trials evaluating drug therapies in RP, in addition to patient-reported outcomes or to establish the proof-of-concept in preliminary studies. However, whether digital perfusion is a valid surrogate for clinical outcomes in RP trials has never been explored. The principal aim of this study was to evaluate the potential surrogacy of digital perfusion, by combining individual-level and trial-level data.

Methods: We used individual data from a series of *n*-of-1 trials, and trial data from a network meta-analysis. We estimated individual-level surrogacy through coefficients of determination between digital perfusion and clinical outcomes (R^2_{ind}). We further calculated the coefficients of determination between treatment effect on the clinical outcomes and on digital perfusion, at the individual level (R^2_{TEind}) and at the trial level (R^2_{trial}), using non-weighted linear regression, with their 95% CI calculated through bootstrapping.

Results: Results from 33 patients and 24 trials were included in the final analysis. At the individual level, there was no correlation between digital perfusion and clinical outcomes at rest and in response to various cooling tests (the highest R^2_{ind} was 0.03 [−0.07, 0.09]), and R^2_{TEind} was also very low 0.07 (0, 0.29). At the trial level, the highest value of R^2_{trial} was 0.1 (0, 0.477).

Conclusions: Digital perfusion, at rest or in response to a cold challenge, and whatever the method used, does not fulfil the criteria of a valid surrogate for existing patient-reported outcomes in RP trials.

Keywords: RP, surrogacy, skin perfusion, SSc

Rheumatology key messages

- Recent trials evaluating drug therapies in Raynaud's phenomenon using patient-reported outcomes have failed to demonstrate treatment efficacy.
- Measurement of digital perfusion, used as an objective outcome, has never been explored as a valid surrogate for patient-reported outcomes in RP trials.
- Our results show that digital perfusion does not fulfil the criteria of a valid surrogate for patient-reported outcomes in RP trials.

Introduction

RP is a clinical condition characterized by transient ischaemia of the extremities in response to cold or emotional stress. It is either primary and benign, or secondary to pathologies such as SSc [1].

Clinical outcomes used in most randomized controlled trials (RCTs) assessing the effect of drug therapies in RP are the frequency of attacks, their cumulative duration and severity scores such as the Raynaud condition score (RCS). However, these patient-reported outcomes (PRO) may be biased due to the subjectivity of self-reporting, and are extremely variable [2] and subject to substantial placebo response [3].

To sidestep these issues, objective outcomes have been widely used in RP trials. They mainly consist in assessing digital perfusion, sometimes in response to a cold challenge. Most of the techniques provide indexes of microvascular blood flow, which is impaired in RP, such as laser Doppler or laser speckle contrast imaging (LSCI) [4], while other methods less specifically measure both the micro- and the microcirculation (such as plethysmography). Some of these tests have shown good convergent validity among each other [5]. However, the surrogacy of these outcomes, i.e. their ability to predict treatment effect on a clinically meaningful outcome, has never been properly evaluated.

Received: 17 April 2023. **Accepted:** 27 June 2023

The validation of a surrogate outcome is typically done at both the individual and the trial levels. At the individual level, a high level of correlation is expected between the clinical outcome and the surrogate outcome; and at the trial level, there must be a good correlation between the treatment effect on the clinical outcome and the treatment effect on the surrogate outcome [6, 7].

The principal aim of this study was to explore whether digital perfusion is a valid surrogate for clinical outcomes in trials assessing the effect of drugs in patients with RP, combining individual-level and trial-level data.

Methods

Data sources

Individual data come from a series of *n*-of-1 trials assessing the efficacy of on-demand sildenafil in RP [8] (ClinicalTrials.gov NCT02050360, approved by the Grenoble Ethics Committee: CPP Sud-EstV). They consist in repeated crossover, double-blind trials during which treatments are administered in repeated blocks. In each block, three treatments were tested: on-demand sildenafil at 40 mg per dose, at 80 mg per dose and placebo. Each patient gave written informed consent before participation.

Trial-level data were collected as an ancillary objective of a systematic review of RCTs, including those using an index of digital perfusion [2]. The full protocol of this review is available online (PROSPERO registry, CRD42017057518). Studies included were double-blind, with parallel or crossover design, they enrolled a majority of patients with SSc-related RP, and two or more pharmacological treatments were tested, or treatment *vs* placebo. Detailed eligibility criteria are available in the supplementary material, available at *Rheumatology* online.

Outcomes

Skin perfusion was assessed on the first day of each treatment period of the first block [8]. A cold provocation test was performed, by placing the right hand during 30 min in a fenestrated box with air at 9 (1)°C. Skin blood flow was measured with LSCI before, during and up to 60 min after cooling, and expressed as the percentage change from baseline in cutaneous vascular conductance [8]. Clinical outcomes were the RCS, the daily frequency and the mean duration of RP attacks at the end of the first block. Individual treatment effect was estimated as the probability of superiority of sildenafil (40 or 80 mg) over placebo, and as adjusted relative variations (aRVs), which represent the magnitude of the effect (i.e. individual effect size) [8].

Trial-level outcomes were indexes of digital perfusion, assessing either the microcirculation only (laser-Doppler, LSCI, thermography), or global digital perfusion, including macrocirculatory blood flow (plethysmography, digital pressure, digital artery ultrasound, temperature to critical closure). They were measured at rest and/or in response to a cold challenge. When there were several indexes of digital perfusion in the same study, they were all extracted. When the same outcome was expressed in different ways, we extracted only the one that was presented first in the results. When indexes of digital perfusion were repeatedly measured during follow-up, we extracted the one which was most concomitant to the clinical outcome. Clinical outcomes were the RCS (or other severity scales), the daily frequency and the mean duration of RP attacks.

Statistical analysis

At the individual level, we assessed the relationship between skin blood flow and PRO using a linear mixed effects model, with fixed effects associated with three treatment categories. In order to meet the condition of application (normally distributed variables), the RCS, the duration of attacks and the daily frequency of RP attacks were transformed using the Box–Cox method. Considering the crossover nature of the original design, each patient was involved in several treatment arms, which is the reason why we added a random intercept modelling the response of each individual. Therefore, parameters of the full model include fixed effects for each treatment category, plus the variance of the patient random effect. We evaluated the strength of these associations with the coefficients of determination (R^2_{ind}), and their 95% CI, computed through bootstrapping.

We further calculated the coefficients of determination between treatment effect on the clinical outcomes and on skin blood flow (expressed as the mean difference between each dosage of sildenafil and placebo) (R^2_{TEind}) with their 95% CI, using the methods described above.

At the trial level, we transformed the different ways of expressing data for the different indexes of digital perfusion into standardized mean differences (SMD) (details are provided in the supplementary material, available at *Rheumatology* online). The strength of the association between treatment effect on indexes of digital perfusion and on clinical outcomes was assessed with the coefficients of determination (R^2_{trial}), using a non-weighted linear regression, with their 95% CI computed through bootstrapping. For studies reporting different indexes of digital perfusion or different comparisons (more than two treatment arms or multiple indexes of digital perfusion), we prioritized laser techniques (Doppler or LSCI), followed by thermography and by global digital perfusion methods.

We considered that surrogacy was reached when the lower limit of the 95% CI of the absolute value of the determination coefficient (R^2) was $\geq$0.72, unclear when the 95% CI lay between 0.5 and 0.72, and absent when the upper limit of the 95% CI was $\leq$0.5, following the German Institute of Quality and Efficiency in Health Care guidelines [9].

All analyses were performed in R version 3.6.1 (R Foundation for Statistical Computing, Vienna, Austria).

Results

Estimation of individual-level surrogacy

Thirty-eight patients were enrolled and 33 completed a full cold challenge for all treatment periods. Among these 33 patients, 11 had RP secondary to SSc.

There was no significant association between skin blood flow at the different times of interest (at rest, during cooling and during rewarming) and the three clinical outcomes (Supplementary Table S1, available at *Rheumatology* online). In Fig. 1, we illustrate the linear regression between skin blood flow and RCS at rest and at the end of 30 min of cooling on the distal phalanx. Results for the other clinical outcomes were comparable to those for RCS. They are provided as Supplementary Figs S1 and S2, available at *Rheumatology* online, for frequency and duration of RP attacks.

Similarly, we found no significant association between treatment effect on skin blood flow and on the three clinical outcomes at the individual level, whatever the dosage

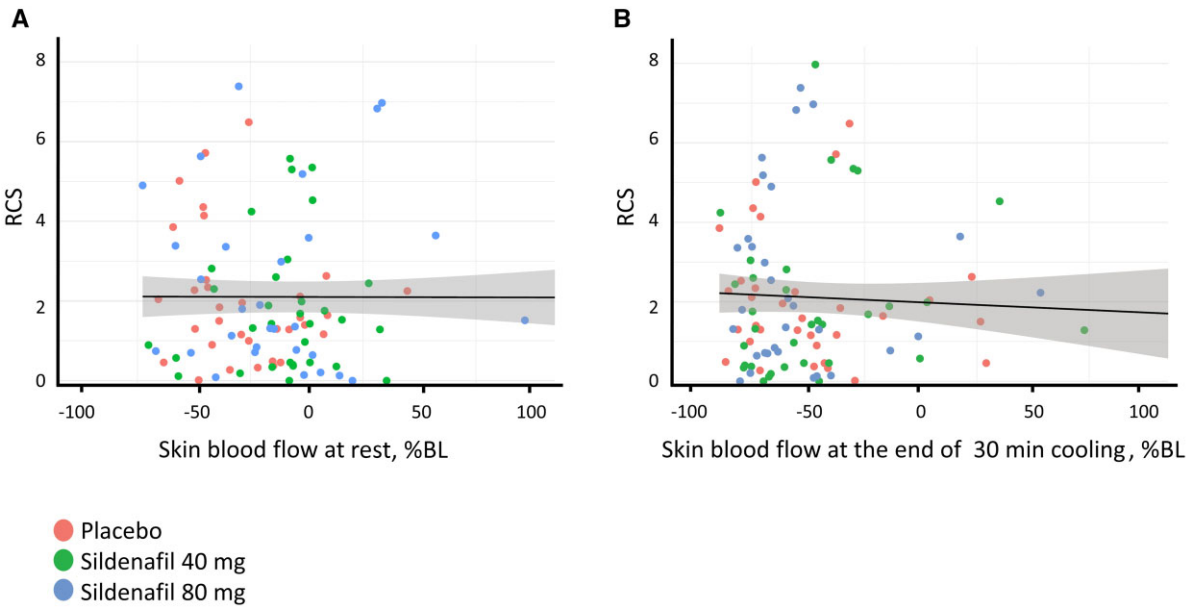

Figure 1. Individual-level correlation between Raynaud's condition score and skin blood flow at rest (**A**) and at the end of 30 min of cooling (**B**) on distal phalanx. Two outliers were removed from the figure for clarity, but kept in the statistical analysis (i.e. to draw the linear regression line). BL: baseline; RCS: Raynaud condition score

(Supplementary Table S2, available at *Rheumatology* online). In Fig. 2A, we illustrated the linear regression between the mean difference of skin blood flow and RCS at rest and at the end of a 30-min cooling on the distal phalanx for sildenafil 40 mg. Full results for the RCS, frequency and duration of RP attacks are represented in Supplementary Figs S3–S5, respectively, available at *Rheumatology* online. Results for sildenafil 80 mg for RCS, frequency and duration of RP attacks are represented in Supplementary Figs S6–S8, respectively, available at *Rheumatology* online.

Estimation of trial-level surrogacy

Final analysis included 78 comparisons from 24 RCTs (listed in the Supplementary Data S1, available at *Rheumatology* online). The flow chart of included RCTs is presented in Supplementary Fig. S9, available at *Rheumatology* online. Among these trials, 13 performed a cold challenge, coupled with different techniques to assess perfusion (Supplementary Table S3, available at *Rheumatology* online). The characteristics of these studies and the description of the methods and temperatures of cold challenges are provided in Supplementary Table S4, available at *Rheumatology* online.

The coefficient of determination R^2_{trial} was 0.002 (0, 0.566), 0.099 (0, 0.477) and 0.009 (0, 0.415) for the RCS, the frequency and the duration of attacks at baseline, respectively, and 0.017 (0, 0.334), 0.000 (0, 0.261) and 0.086 (0, 0.667) for the same outcomes after a cold challenge. Linear regressions between treatment effect on digital perfusion and treatment effect on each of the three outcomes are shown in Fig. 2B and Supplementary Figs S10–S12, available at *Rheumatology* online.

Discussion

This study is the first, to our knowledge, to assess the validity of digital perfusion as a possible surrogate in RP trials. Our results clearly demonstrate the lack of association between objective and subjective outcomes (at the individual level) and the lack of association between treatment effect on digital perfusion and on clinical outcomes, both at the individual and at the trial levels.

The rationale for using digital perfusion as a surrogate is based on the pathophysiology of RP and the pharmacology of most drugs, which are vasodilators that aim at maintaining adequate perfusion in the fingers. Among them, phosphodiesterase type 5 inhibitors (iPDE5) have shown significant (although modest) efficacy on clinical outcomes [2], as well as on digital perfusion at rest and during or after cooling [8]. These two conditions correspond to the first and second criteria required to validate a surrogate according to Prentice. Yet, it is not sufficient to use an outcome as a surrogate. Indeed, both clinical and surrogate outcomes must correlate, and importantly, treatment effect on the surrogate must reasonably predict treatment effect on the clinical outcome.

At the individual level, there was no significant association between skin perfusion and clinical outcomes, or between treatment effect of sildenafil on these outcomes and on skin perfusion at rest and in response to the cooling test. Indeed, our results show particularly poor fit of the model with fixed effects, thus suggesting that all the variance is explained by the random factor, i.e. the patient. Similarly, our results do not show a strong association between treatment effect on indexes of digital perfusion and on clinical outcomes, i.e. indexes of digital skin blood flow used in RP trials do not predict treatment effect on PRO. It is worth noting that there is no consensus on the value of the threshold used to define surrogacy. Although we chose a rather strict threshold, the coefficients of determination we observed were far below acceptable values, and the choice of lower threshold [10] would have not changed the conclusions.

At the trial level, the heterogeneity in methods used to assess digital perfusion is a limitation which may explain the poor matching. There are different types of cooling tests, and a variety of temperatures, medium and devices. Similarly, the

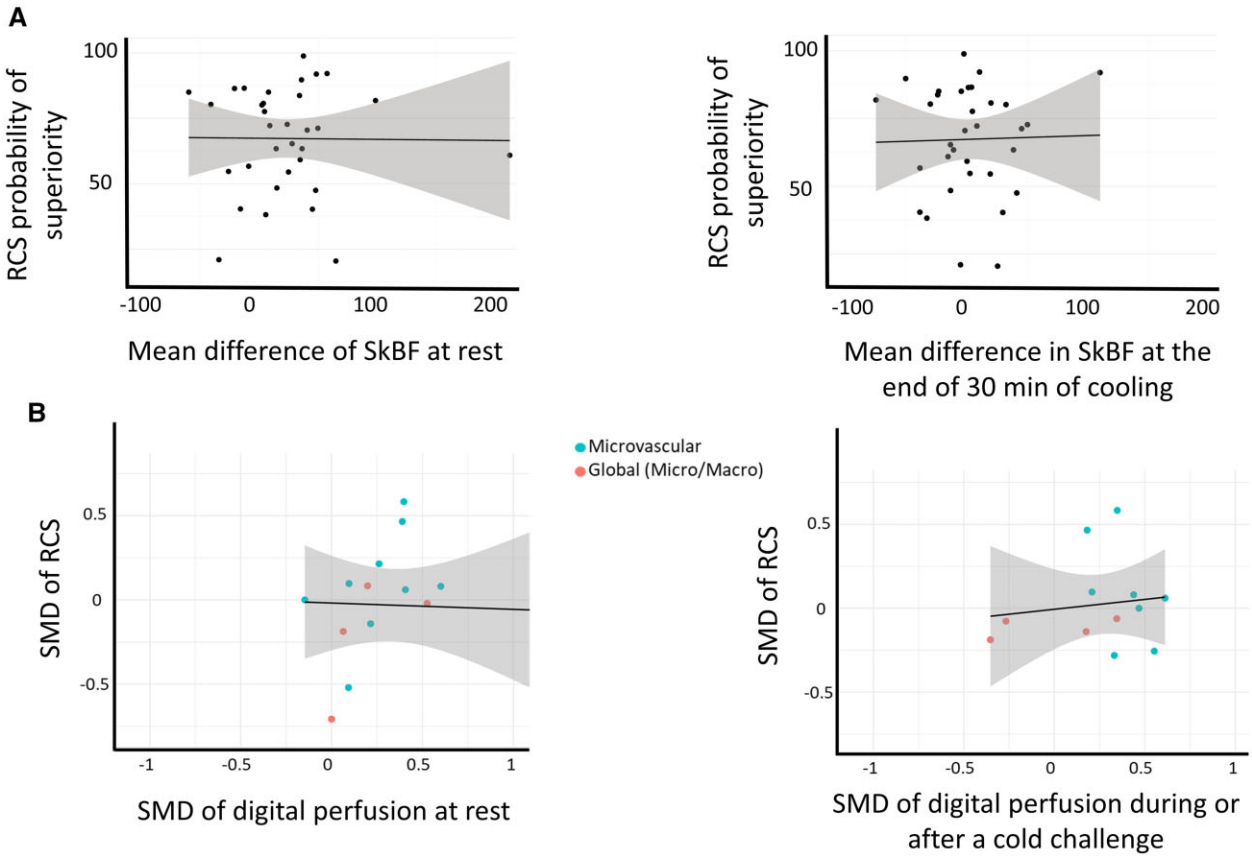

Figure 2. Regression analysis between the treatment effect on Raynaud's condition score and on skin blood flow at individual level (**A**) or digital perfusion at trial level (**B**) at rest and during or after a cold challenge. Individual data are expressed as mean difference between skin blood flow with treatment (sildenafil 40 mg) and skin blood flow with placebo, and the probability of superiority *vs* placebo at rest and at the end of 30 min of cooling. Trial-level data are expressed as standardized mean differences (SMD) at rest and during or after a cold challenge. RCS: Raynaud condition score; SkBF: skin blood flow

multiplicity of techniques to assess perfusion and differences in the expression of the results add variability. In addition, assessment of digital perfusion is sensitive to the experimental conditions (temperature, humidity) and the patient's state (duration of the acclimation and resting period, fasting or not, tobacco or coffee consumption, etc.). While room temperature was controlled for digital perfusion measurements, and daily temperature included in the model estimating treatment effect on PRO at the individual level, this information is lacking at the trial level. A recent review suggests that laser-derived methods to measure skin blood flow could potentially be a surrogate for SSc-related digital vasculopathy. Yet, this qualitative synthesis suggests that their association with outcomes is greater for treatment effect on digital ulcers than on RP [11], which is consistent with our findings. In that review, full-field techniques have been suggested to be superior to other methods [11], and LSCI has indeed been widely used to assess skin microvascular function in SSc and RP over the past few years [4, 5, 8, 12]. Yet, the results from our series of *n*-of-1 trials, in which we have used a fenestrated cooling box with continuous recording using LSCI, do not support that LSCI is a valid surrogate for PRO in RP.

Limitations

We included a relatively limited number of studies, which explains the large confidence intervals of the coefficients of determination. However, point estimates of R^2 are so low that it seems unlikely that narrower confidence intervals would allow reaching surrogacy criteria. Another limitation is that we included data from patients with primary and secondary RP, which have different pathophysiological features. Therefore, we cannot exclude a difference in surrogacy between these two populations. However, the small sample size of this study does not permit us to explore whether this difference exists. Similarly, we included a wide range of dosages and forms in the trial-level analysis. Yet, such heterogeneity is not an issue in the validation of a surrogate outcome, since the latter should fully capture the treatment effect on the clinical outcome, whatever its magnitude.

Another limitation of our work is that among existing PROs used in RP trials, the most common, i.e. the RCS, has only been validated in patients with SSc-related RP. Moreover, it has never resulted in positive trial outcomes. There is ongoing work to develop PROs that more accurately capture the burden of RP on patients' function and daily life [13]. Additional work will be needed to evaluate the surrogacy of objective measures of digital perfusion on this outcome.

Future studies should be conducted to assess whether other methods (such as a portable device) are more reliable tools to predict treatment effects. The relationship between these new devices and questionnaire results deserves to be assessed in future work [14].

Finally, the absence of valid surrogacy does not mean that indexes of digital skin perfusion are useless in RP trials. Although, they are not sufficient to conclude on the efficacy

of therapies, they can be used as exploratory endpoints in early-phase, proof-of-concept studies [15].

In conclusion, this study shows that treatment effects on clinical outcomes are not strongly associated with treatment effects on digital skin perfusion in RP trials, either at the individual or at the trial level. Therefore, digital skin perfusion at rest or after a cold challenge is not a valid surrogate for existing patient-reported outcomes in RP.

Supplementary material

Supplementary material is available at *Rheumatology* online.

Data availability

The full dataset used for individual-level analyses is available at https://datadryad.org/resource/doi:10.5061/dryad.c670tq2. The full dataset used for trial-level analyses is available at https://osf.io/wxf6m.

Contributiion statement

Conception of the work: CK, MR; acquisition of the data: AG, LL, OG, CK, MR; data analysis: AG, LL, MM, CK; interpretation of the data: MM, JLC, SB, CK, MR; drafting the manuscript: AG, MR; revision of the manuscript for important intellectual content: LL, MM, OG, JLC, SB, CK; final approval of the version to be published: all authors.

Funding

This work has been partially supported by MIAI @ Grenoble Alpes (ANR-19-P3IA-0003).

Disclosure statement: The authors have declared no conflicts of interest.

References

1. Herrick AL, Wigley FM. Raynaud's phenomenon. Best Pract Res Clin Rheumatol 2020;34:101474.
2. Khouri C, Lepelley M, Bailly S *et al.* Comparative efficacy and safety of treatments for secondary Raynaud's phenomenon: a systematic review and network meta-analysis of randomised trials. Lancet Rheumatol 2019;1:e237–46.
3. Roustit M, Jullien A, Jambon-Barbara C *et al.* Placebo response in Raynaud's Phenomenon clinical trials: the prominent role of regression towards the mean: placebo response in Raynaud's Phenomenon. Semin Arthritis Rheum 2022;57:152087.
4. Ruaro B, Sulli A, Alessandri E *et al.* Laser speckle contrast analysis: a new method to evaluate peripheral blood perfusion in systemic sclerosis patients. Ann Rheum Dis 2014;73:1181–5.
5. Wilkinson JD, Leggett SA, Marjanovic EJ *et al.* A multicenter study of the validity and reliability of responses to hand cold challenge as measured by laser speckle contrast imaging and thermography: outcome measures for systemic sclerosis-related Raynaud's Phenomenon. Arthritis Rheumatol 2018;70:903–11.
6. Buyse M, Molenberghs G, Paoletti X *et al.* Statistical evaluation of surrogate endpoints with examples from cancer clinical trials. Biom J Biom Z 2016;58:104–32.
7. Alonso A, Molenberghs G, Geys H, Buyse M, Vangeneugden T. A unifying approach for surrogate marker validation based on Prentice's criteria. Stat Med 2006;25:205–21.
8. Roustit M, Giai J, Gaget O *et al.* On-demand sildenafil as a treatment for Raynaud phenomenon: a series of *n*-of-1 trials. Ann Intern Med 2018;169:694–703.
9. Institute for Quality and Efficiency in Health Care: Executive Summaries. Validity of surrogate endpoints in oncology executive summary of rapid report A10-05, Version 1.1. Cologne, Germany: Institute for Quality and Efficiency in Health Care (IQWiG), 2005. http://www.ncbi.nlm.nih.gov/books/NBK198799/ (16 August 2022, date last accessed).
10. Ciani O, Buyse M, Drummond M *et al.* Use of surrogate end points in healthcare policy: a proposal for adoption of a validation framework. Nat Rev Drug Discov 2016;15:516.
11. Pauling JD, Hackett N, Guida A, Merkel PA. Performance of laser-derived imaging for assessing digital perfusion in clinical trials of systemic sclerosis-related digital vasculopathy: a systematic literature review. Semin Arthritis Rheum 2020;50:1114–30.
12. Guigui A, Mazet R, Blaise S *et al.* Treprostinil hydrogel iontophoresis in systemic sclerosis-related digital skin ulcers: a safety study. J Clin Pharmacol 2020;60:758–67.
13. Yu L, Domsic RT, Saketkoo LA *et al.* The Assessment of Systemic sclerosis-associated RAynaud's Phenomenon (ASRAP) questionnaire: item Bank and Short Form Development. Arthritis Care Res 2023;75:1725–34.
14. Herrick AL, Heal C, Wilkinson J *et al.* Temperature response to cold challenge and mobile phone thermography as outcome measures for systemic sclerosis-related Raynaud's phenomenon. Scand J Rheumatol 2021;50:479–84.
15. Herrick AL, Batta R, Overbeck K *et al.* A Phase 2 trial investigating the effects of the angiotensin II type 2 receptor agonist C21 in systemic sclerosis-related Raynaud's. Rheumatology 2023;62:824–8.