Foeldvari Ivan (Orcid ID: 0000-0003-0659-5298)
Klotsche Jens (Orcid ID: 0000-0002-2954-5755)
Anton Jordi (Orcid ID: 0000-0002-8792-4219)
Del Galdo Francesco (Orcid ID: 0000-0002-8528-2283)
Denton Christopher P (Orcid ID: 0000-0003-3975-8938)
Cutolo Maurizio (Orcid ID: 0000-0002-5396-0932)
Kasapcopur Ozgur (Orcid ID: 0000-0002-1125-7720)
Khanna Dinesh (Orcid ID: 0000-0003-1412-4453)
Ruperto Nicolino (Orcid ID: 0000-0001-8407-7782)
Furst Daniel E. (Orcid ID: 0000-0001-7857-2373)

Title: Proposed response parameters for 12 months drug trial in Juvenile Systemic Sclerosis. Results of the Hamburg International Consensus Meetings

Authors:

Ivan Foeldvari* MD[1], Kathryn S. Torok* MD[2], Jens Klotsche Dr.rer.nat.[3], Jordi Anton MD[4], Patricia Cost Reis MD[5], Francesco Del Gado MD[6], Christopher Denton MD[7], Bernd Hinrichs MD[8], Suzanne C. Li MD[9], Clare Pain MD[10], Clarissa Pilkington MD[11], Maurizio Cutolo MD[12], Ozgur Kasapcopur MD[13], Dinesh Khanna MD[14], Nicola Ruperto MD[15], Vanessa Smith MD[16], and Daniel E. Furst MD[17]

*equal first authors

Affiliations:

1 Hamburg Centre for Pediatric and Adolescent Rheumatology, Schön Klinik Hamburg Eilbek, Hamburg, Germany; 2 University of Pittsburgh, Children's Hospital of Pittsburgh, Pittsburgh, PA, USA; 3 German Rheumatism Research Center, Berlin, Germany; 4 Pediatric Rheumatology Division. Hospital Sant Joan de Déu Barcelona, Universitat de Barcelona, Barcelona, Spain; 5 Pediatrics Department, Hospital de Santa Maria, Faculdade de Medicina, Universidade de Lisboa, Lisbon, Portugal; 6 Institute of Rheumatic and Musculoskeletal Medicine, University of Leeds, Leeds Teaching Hospital Trust, Leeds, UK; 7 UCL Division of Medicine, Royal Free London NHS Foundation Trust, London, UK; 8 Children's pulmonology, Asklepios Klinik Nord – Heidberg, Hamburg, Germany; 9 Hackensack University Medical Center, Hackensack, NJ, USA; 10 Alder Hey Children's Foundation NHS Trust, Liverpool, UK; 11 Great Ormond Street Hospital, London, UK; 12

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/acr.25171

University of Genoa, Genoa Italy; 13 Department of Pediatric Rheumatology, Cerrahpasa Medical School, Istanbul University - Cerrahpasa, Istanbul, Turkey; 14 University of Michigan Scleroderma Program, Division of Rheumatology/Dept. of Internal Medicine, Ann Abor, Michigan, USA; 15 IRCCSG Istituto G. Gaslini, Genoa, Italy; 16 Department of Internal Medicine, Ghent University, Ghent, Belgium; Department of Rheumatology, Ghent University Hospital, Ghent, Belgium; Unit for Molecular Immunology and Inflammation, VIB Inflammation Research Center (IRC), Ghent, Belgium, Member of ERN ReCONNET (rare connective tissue diseases), Ghent, Belgium; 17 University of California, Los Angeles; University of Washington, Seattle, Washington; University of Florence, Florence Italy

Corresponding Authors:

Ivan Foeldvari, MD

Hamburg Centre for Pediatric and Adolescence Rheumatology

Centre for Treatment of Scleroderma and Uveitis in Childhood and Adolescence

Teaching Unit of the Asklepios Campus of the Semmelweis Medical School, Budapest

An der Schön Klinik Hamburg Eilbek

Dehnhaide 120

22081 Hamburg

Tel: +49 40 2092 3697

Fax: +49 40 2092 3693

Email: sprechstunde@kinderrheumatologie.de

Kathryn S. Torok, MD

Associate Professor of Pediatrics

Division of Pediatric Rheumatology

UPMC & University of Pittsburgh Scleroderma Center

Director, Pediatric Scleroderma Clinic

University of Pittsburgh |UPMC Children's Hospital of Pittsburgh

4401 Penn Ave

Pittsburgh, PA 15224

Tel: (412) 692-5081

Fax: (412) 692-5054

Email: kathryn.torok@chp.edu

Word Count: 4019

Tables: 1

Figures: 1

Supplementary tables: 1

Supplementary materials: 2 (Physician facing outcomes and Patient facing outcomes clinical research forms)

Financial support: No financial support or other benefits from commercial sources were obtained for the work reported on in the manuscripts. There is no financial conflict of interest of any of the authors.

Objectives:

Juvenile systemic sclerosis (jSSc) is an orphan disease, associated with high morbidity and mortality. New treatment strategies are much needed, but it is necessary to clearly define appropriate outcomes if successful therapies are to be developed. Here, such outcomes are proposed.

Methods:

This proposal is the result of four face to face consensus meetings with a 27-member multidisciplinary team of pediatric rheumatologists, adult rheumatologists, dermatologists, pediatric cardiologists, pulmonologists, gastroenterologists, statistician and patients. Throughout the process, we reviewed the existing adult data in this field, the more limited pediatric literature for jSSc outcomes and data from two jSSc patient cohorts to assist in making informed, data-driven decisions. The use of items for each domain as an outcome measure in an open 12-month clinical trial of jSSc was voted and agreed upon using a nominal group technique.

Results:

After voting, the agreed domains were: global disease activity, skin, Raynaud's phenomenon, digital ulcers, musculoskeletal, cardiac, pulmonary, renal, gastrointestinal, and quality of life. Fourteen outcome measures had 100% agreement, one item had 91% agreement and one item had 86% agreement. The domains of biomarker and growth/development were moved to the research agenda.

Conclusion:

We reached consensus on multiple domains and items which should be assessed in an open label 12-month clinical jSSc trial as well as a research agenda for future development.

Significance and Innovations

- This is the first proposal for outcome measures for a 12-month clinical trial for jSSc
- The proposed outcome measures span the main domains of the organ system involvement in jSSc
- Patient reported outcomes are included in the outcomes proposed

Introduction

Juvenile systemic sclerosis (jSSc) is an orphan disease with an estimated prevalence of 3 in 1,000,000 children with a high morbidity (1, 2). Currently no medications are licensed for jSSc. This proposal is for an open label, 12-month clinical trial in jSSc. To develop such a trial, and for use in any well-done treatment trial in jSSc, it is necessary to clearly define outcomes and to tailor the outcomes for jSSc.

The only existing outcome scoring system that was developed specifically for use in the jSSc population is the Juvenile Systemic Sclerosis Severity Score (J4S), which was adapted from an adult SSc severity score, the Medsger Severity Score (3, 4). More recent efforts have been proposed for disease activity indices in SSc (rather than severity/damage scores), which would be more sensitive to clinical change and applicable to clinical trials. This includes the Composite Response Index in Systemic Sclerosis (CRISS) (5) which is based on a two-step approach. First, significant disease worsening or new-onset organ damage is defined as non-responsiveness. In patients who did not fulfill the criteria of part one, a probability of improvement is calculated for each patient based the Rodnan Skin Score (mRSS), percent predicted forced vital capacity (FVC%), physician global assessments (PGA), and the patients' Health Assessment Questionnaire Disability Index (HAQ-DI). These efforts applied only to adult SSc (5, 6, 7) and there is no such disease activity or response index for jSSc. Our goal was to define disease activity outcome parameters in jSSc, which would be sensitive to change and useful in an open label, 12-month clinical trial in jSSc.

Methods

These recommendations were developed over a span of years by a dedicated group of multinational pediatric and adult scleroderma experts who are interested in jSSc and outcome measure development, starting with electronic surveys in 2014 and refining jSSc outcome domains and items through an annual face-to-face meeting, through Delphi and Nominal group technique processes, hosted at the Hamburg Symposium of Juvenile Scleroderma starting in 2014 – 2018 (see **Figure 1** for details). The final 2018 jSSc

consensus meeting is explained in detail here. By consensus, it recommended 12 domains and 22 items for an open-label, 12-month clinical jSSc trial.

In December 2018, international pediatric and adult rheumatology scleroderma experts, dermatologists, pediatric cardiologists, pulmonologists, gastroenterologists, statistician and patients, met for a two-day conference. The first day was dedicated to a series of talks and discussions regarding adult scleroderma expert presentation of the CRISS (DK) and items included in the CRISS with possible pediatric performance and adaptations(8) , a 'lessons learned' talk regarding response of clinical outcomes from recent clinical trials in adult SSc (CD), cohort data from the International juvenile systemic sclerosis Inceptions cohort (jSSc Inceptions cohort; n=150)(9) and the Childhood Arthritis and Rheumatology Research Alliance jSSc cohort (CARRA jSSc cohort; n=64)(10) relevant to the prior voted domains and items of interest, and finally, by pediatric scleroderma workgroup presentations on updates of the various organ systems in jSSc and related outcomes. These discussions provided background for the second day of the conference, whose goal, using the nominal group technique (DEF), was to develop consensus recommendations for items to be used in an open, 12-month clinical trial in children with systemic sclerosis (not clinical practice or general research). The items (n=22) and domains (n=12) remaining after the 2017 Hamburg consensus meeting (**Fig. 1**) were re-evaluated at the 2018 meeting. Twenty-two out of the 27 multidisciplinary members at the 2018 conference voted, with 75% (16/22) having been at the preceding 2017 consensus conference (**Supplemental Table 1**).

General Guidelines

Some general guidelines were discussed and agreed at the start of the second day of the consensus 2018 meeting, including the following: validated outcomes should have priority; outcomes validated in adults with SSc would be sufficient for application in jSSc; although previously agreed items (from the 2016 and the 2017 consensus meetings) were defined in terms of change, those items will be defined in terms of their absolute value, independent of change per se (e.g. the item "change in mRSS", would now be "mRSS"); and estimation of change from baseline and significance of change would be examined through statistical analysis. For uniformity and clarity when patients or clinicians used the measures, a scale of 0-10 or 0-100

was to be employed when visual analogue scales (VAS) were used as items. The specific length of the scale could be decided on a protocol basis. The MCID for any VAS was to be 1.0 for 0-10 and 10.0 for 0-100 scales. This decision was slightly less than that in the literature (7-27 mm dependent on baseline pain) but was felt easy to use and remember (22/22 agreed). It should be noted that, unfortunately, other MCIDs were usually not available for consideration or voting. There was also consensus (22/22 agreed) that the time frame for any VAS was to be 7 days unless specifically stated differently. The CRISS, a validated combined measure of response in adult SSc, although discussed at length the day prior, was not voted upon as a composite outcome during the second day consensus meeting since it comprised multiple important elements which were instead individually voted on within their respective domain.

The Consensus Process

The process included the following: review of each of the 12 domains and items within each domain, led by the moderator (DEF). There was first to be some minimal background given for orientation (usually from the leader of the organ working group); during discussion, there was to be one speaker at a time, voting (22 members, later 21 members as one member had to leave) would close the discussion and consensus (**Supplemental Table 1**). Voting was not anonymous, and options included agree, disagree or do not know. Consensus was defined a priori, as agreement by >= 80% of voting members present. If consensus was not reached, more discussion ensued, and, ultimately, without consensus, the item was recorded as "no consensus reached", and, if applicable, referred to the research agenda.

There were three scribes (KT, MC, NH), who compared notes after the meeting to ensure accuracy. KT merged notes and DEF reviewed and edited. There were invited consultants who participated in the prior day's meeting and provided some discussion points during the consensus meeting, but refrained from voting (BH - pulmonary, LA – gastrointestinal, CB – cardiology) (**Supplemental Table 1**). There were two SSc patients present (AZ and KF), both currently adults with ages of onset of 8 years old and 26 years old, who actively participated and were voting members.

Results

Domain 1. Patient and Physician Global Disease Activity

The physician's and the patient's VAS (0-100 mm) global assessment of disease activity (PGA-A and PtGA-A) over the previous 7 days have been used in the jSSc Inceptions cohort (1), with data in 47 jSSc patients over 12 months demonstrating an MCID of 20/100 mm change ($p<0.001$) in physician global and a 15/100 change in patient global ($p<0.001$). **There were unanimous votes (22/22 for each) to use the PGA-A and PtGA-A in jSSc trials (Table 1; Supplemental Clinical Research Forms (CRFs)).** There was consensus to include knowledge of the patient's previously available clinical data (19/22 agreed, 3/22 disagreed, 0/22 did not know). By general agreement, it was recommended that there was to be instructions in the protocol or in the "Manual of Procedures" as to how the PGA-A was to be done and be performed by the same investigator at each visit. It was also agreed by all that either child or parent may answer the PtGA-A (age and child dependent) so long as it is consistent throughout the protocol. Patients ages 8 years and older are encouraged to complete patient reported outcomes (PROs) as routine for several pediatric rheumatology registry studies(10).

Domain 2. Patient Reported Global Health and Function

PROs are essential in clinical drug trials. For jSSc, several PROs, including QoL measures, were voted in 2017 and re-supported in 2018. The PRO measures to include were unanimously agreed (22/22 agreed) to be the Childhood Health Assessment Questionnaire (CHAQ) (11, 12) and the Scleroderma-specific visual analogue scales (VAS), derived from the Scleroderma-HAQ Disability Index (S-HAQ-DI)(13, 14) (**Table 1**). **Supplemental CRF**, patient-facing, provide detail of questions.

The CHAQ is a standard PRO, a child-directed assessment of function, modified from the multiply validated adult HAQ-DI (11, 12, 13, 14, 15, 16) The CHAQ ascertains results over eight functional domains and it has been used in two large jSSc cohorts (international inception and CARRA jSSc cohorts, respectively) with mean

scores of 0.45 and 0.40 (range 0-3) (1, 10). The CHAQ, although it has floor effects, reflects the domains which are important to the function of the patient. Thus, it correlates with Global well-being, HRQOL and organ systems of importance to patients with jSSc (10). **The group voted unanimously (22/22) in favor of including the CHAQ for SSc patients less than 16 years old and the HAQ-DI for patients 16 and older.**

No formal MCID has been defined for the CHAQ in jSSc but the jSSc inceptions and the CARRA network cohorts demonstrated that jSSc patients improved by 24 and 44% over 1 and 2 years (respectively), and corresponded with improvements in PtGA-A and PGA-A (P=0.02) ((1),23,(10)). Voting at the consensus meeting then took place in regard to MCID of CHAQ in jSSc and it was agreed to apply the MCID and cut-points from JIA and adult RA to the jSSc cohorts (20/22 agreed, 0/22 disagreed, 2/22 did not know). For reference, among 67 JIA patients followed longitudinally, those rated without change had median CHAQ change of 0, and for those rated as having change, the MCID was -0.188 for improvement and +0.125 for worsening (17).

The other main group of PRO measures discussed were VASs captured in the Scleroderma HAQ (S-HAQ-DI). In jSSc, these VAS scales have been piloted in the National Registry of Childhood Onset Scleroderma cohort (NRCOS; PI – Torok) (n = 20) and were one of the few to have direct patient input (unpublished). **All of the VAS captured in the Scleroderma-HAQ-DI (S-HAQ-DI), were voted upon in their respective organ or general categories and it was determined that they should be included by a unanimous 22/22 vote as important PRO measures in a jSSc trial (Table 1)**. These include the following: Pain overall, Gastro-Intestinal problems, Breathing problems, Raynaud's severity, Finger ulcer severity and Patient global, which capture the SSc patient's perspective on the level of interference with normal activity in these domains over the past week (13). Modifications to the questionnaire for patients < 8 years of age may say "your child" instead of "you".

Further discussion regarded the numerous other pediatric QoL instruments available and validated in other connective tissue diseases, particularly JIA. Available QoL instruments include: Peds QL (18), Peds Rheum QL

(19), Family QL (20), CHQ (21), and CHU-9D (22). These measures were not included in the current published jSSc cohort and so their performance characteristics are unknown in jSSc. Although which QoL instrument is to be used is unknown **it was decided unanimously (22/22) that QoL, (in addition to the CHAQ and S-HAQ-DI VASs) are important to capture in jSSc patients**.

Domain 3. Skin

The modified Rodnan Skin Score (mRSS) is a pivotal outcome measure for any therapeutic trial in diffuse adult and juvenile-onset SSc, assessing the degree of skin thickness over 17 body sites (0 - 3 per skin area, range: 0-51). The mRSS is fully validated via OMERACT filters in adult SSc (23). The mRSS has not been formally validated in jSSc; however, it has been widely adopted in clinical practice and larger observational cohort studies (1, 24). The **mRSS was unanimously agreed at the 2017 and 2018 (22/22 agreed) meetings as the only item in the skin domain.**

Since the mRSS was developed, studied and validated in adult SSc patients, with typical average age of onset between 40 to 50 years old, a few cutaneous variables to consider in the scoring approach in children were suggested (based on expert opinion). These pediatric rheumatology experts considered additional qualitative features, such as the texture of the skin (i.e. waxy, smooth, hard) compared to other areas in that region of the body, the appearance of the skin (i.e. shininess, yellow/waxy appearance), lack of hair, thin skin with visible veins, dyspigmentation and atrophy (dermal or subcutis)(2).

While to be used in jSSc clinical trials, mRSS needs further thorough examination in jSSc in the future (25). As no MCID has been developed**, it was decided to use the absolute mRSS and a statistical change as a measure of skin response in a jSSc trial (22/22 agreement) (Table 1).**

Domain 4. Raynaud's Phenomenon

SSc-associated Raynaud's Phenomenon (SSc-RP) is the most common disease-specific manifestation of SSc (26). SSc-RP was ranked by adult patients as having the highest impact on QoL and perception of illness severity (27). RP was recorded in 75% of the patients in the jSSc inception cohort (28). In a clinical trial, RP should be measured in a standardized manner to assess whether a proposed new treatment is effective. Raynaud's outcomes are primarily patient reported, including frequency, severity and duration but may be confounded by pain and coping strategies(26, 29) .

In the 2017 jSSc meeting, 24 of the 25 participants voted that RP should be assessed. After some discussion among Raynaud's Condition score, Raynaud's VAS from the scleroderma HAQ-DI, and the Physician's assessment of Raynaud's, **the Raynaud's VAS from the scleroderma HAQ-DI was agreed for a jSSc trial (22/22) (Table 1).**

As no MCID was available, it was agreed (22/22) to **use a statistically significant difference in the VAS across timepoints as a useful measure in jSSc trials.**

Domain 5. Digital Ulceration

SSc related digital ulcers (DU) are a frequent and disabling clinical complication of jSSc, affecting approximately 50% of patients in the cohort of 150 patients(1). DU occur most frequently on the fingers or toes and can be the consequence of endothelial damage, trauma, or calcinosis. DU impair hand function and compromise patients' QoL (30). To measure the burden of finger/digital/skin ulcers, the DU clinical assessment score (DUCAS) was developed and validated in adult SSc patients (31). The DUCAS captures the number of DU, new DU, gangrene, surgery needed, infection, unscheduled hospitalization for DU, analgesics for DU pain (most in a yes/no fashion). The DUCAS plus the digital ulcer Scleroderma-HAQ VAS, encompass the items suggested in a survey of the EUSTAR regarding DU impact in SSc (32).

It was unanimously voted (22/22) to include the DUCAS score and the as digital ulcer Scleroderma-HAQ VAS as outcomes measure for digital ulcers in a jSSc trial (Table 1).

Domain 6. Musculoskeletal System

Musculoskeletal manifestations, including joint, muscle and/or tendon involvement occur in 75-82% of jSSc patients, with 19% having documented inflammatory arthritis in prospective cohort studies (1, 10, 24). In 2017, several variables constituting musculoskeletal involvement were considered, including swollen joint count, limited joint range of motion, change in muscle strength assessed by childhood myositis assessment scale or manual muscle testing, new occurrence of tendon friction rubs, and change in muscle enzyme levels (creatine kinase, aldolase). The group reached consensus agreement on including swollen joint count and not the other discussed variables.

Although swollen joints were agreed to be included in jSSc trial in 2017, this was again discussed in 2018, pointing out that measuring only swollen joints captured only a portion of the musculoskeletal involvement in jSSc (e.g. not capturing tenosynovitis, contracture), while also missing inflammation because joint swelling is difficult to measure in SSc (33). Ultimately, **it was decided by a unanimous vote (22/22) to collect musculoskeletal involvement in a jSSc trial as the number of joints that have either joint swelling or limitation in range of motion associated with joint pain or tenderness that is considered secondary to jSSc, thus including tenosynovitis (Table 1)**. It will be called the '**active joint count'**, will be a total score and will be very similar to the joint counts in JIA (thus not requiring special training).

Domain 7. Cardiac Involvement

Although in jSSc, cardiac involvement is relatively infrequently clinically detected (5-15%), it is one of the major causes of non-infectious mortality in jSSc (24, 34). A consensus meeting among European cardiologists and rheumatologists(35) indicated the need to examine for arrhythmias (EKG, Holter monitor), an imaging

measure to examine fibrosis (i.e MRI of the heart), plus patient response outcomes, and echocardiogram to define cardiac involvement in SSc.

Several cardiac variables were discussed in the context of the 2017 and 2018 jSSc meeting and there was 100% agreement (22/22) on the following parameters (**Table 1**):

(1) a measure of ejection fraction was appropriate as an inclusion measure

(2) new onset of left ventricular failure and/or new "clinically important arrhythmia (malignant/non-malign)" were appropriate measures defining lack of response in a jSSc trial

(3) the development of pulmonary hypertension 'by accepted criteria' is a sign of nonresponse

(4) the development of new-carditis should be removed from consideration (not well defined)

(5) the NT-proBNP, not validated in jSSc, was to go into the research agenda

Noted was that two of these consensus items are included as the Step 1 CRISS criteria for adult SSc: new onset left ventricular failure and new onset pulmonary hypertension, though both are specified further in adult SSc with "≤45% ejection fraction requiring treatment" and "measured via right cardiac catheterization requiring treatment", respectively (5).

Domain 8. Pulmonary Involvement

Interstitial lung disease (ILD) occurs in approximately 50% of patients in jSSc (1). It is a major reason for mortality in adult patients with SSc (5, 36). Screening for ILD in adult and pediatric SSc patients traditionally includes pulmonary function testing (PFT) with forced vital capacity (FVC) and single breath diffusion capacity for carbon monoxide (DLC0) (37, 38). In children, assessment of FVC is fairly standardized from age 3 years, while DLCO is more reliable starting at age 8 years (39, 40). The combination of high resolution computed tomography (HRCT)(low radiation protocols) and PFTs are now used to both detect and follow ILD progression and regression in adults (37, 41). In children, HRCT has been eschewed because there is concern regarding radiation (39, 40). The 6 Minute Walk Test (6MWT) is a sensitive measure with an MCID of 10

meters (42, 43) and normal values for healthy children exist for comparison (44). Although in adults with SSc the 6MWT is not responsive to treatment as it is confounded by joint contracture, muscle weakness and fatigue.

At the 2017 consensus group meeting it was agreed that the core CRISS variables, including the change in FVC (5), were appropriate for jSSc and in the 2018 consensus meeting there was 100% consensus **to include FVC and age eligible DLCO in jSSc trials.** The group decided to **include the 6MWT assessment in the core set** (18/21 agreed, 3/21 disagreed), measured as absolute meters with within patient changes for statistical comparisons (**Table 1**).

Because there remained concerns of increased risk of malignancy after repeated HRCT of the lungs (45), the group unanimously rejected it as a required outcome measure in a jSSc trial.

Domain 9. Renal Involvement

The course of renal involvement in jSSc is usually benign but there is a broad spectrum of renal manifestations in jSSc, from mild proteinuria to acute renal failure. The most severe type is characterized by new onset hypertension accompanied by acute kidney injury, proteinuria, hematuria or signs of microangiopathy (thrombocytopenia or hemolysis) (Scleroderma renal crisis; SRC). SRC is a rare event in children(1), but it remains a major risk factor for mortality.

The consensus group agreed unanimously (21/21) to **include the new occurrence of scleroderma renal crisis as an outcome measure criterion for a jSSc trial (Table 1)**. This is also an adult CRISS Step 1 criterion, which would consider the patient as not improved (5). It was adjusted, accounting for the definition of high blood pressure in children and adolescents (46) and the Kidney Disease: Improving Global Outcomes (KDIGO) definition of acute kidney injury (47).

Other items related to renal involvement, namely new diagnosis of hypertension, new persistent proteinuria and decrease on glomerular filtration rate, were unanimously (21/21) rejected by the group, as outcome

measure criteria for jSSc treatment trials, because they lacked specificity and/or had a low prevalence in jSSc patients.

Domain 10. Gastrointestinal Involvement

Gastrointestinal (GI) manifestations of SSc have been reported in 25-92% of children and are associated with poor quality of life (10, 48). GI manifestations in adult SSc patients range from mild oropharyngeal dysphagia to malnutrition (15-56%) and increased mortality (49). Malnutrition is a major concern in the growing child and has been shown to predict mortality in other pediatric chronic illnesses with GI absorption issues, such as chronic kidney disease(50). Typical measures in children to assess malnutrition include midarm circumference (MAC) and triceps skinfold (TSF) thickness(51) however, in jSSc these measures may be confounded by skin manifestations. Another indicator of malnutrition in children, very low Body Mass Index (BMI) (z-score $\leq$ 2), indicating moderate to severe malnutrition, can be used in jSSc, with very low BMI documented in 14% of the jSSc CARRA registry patients and correlating with poor QoL measures(10). Multiple other non-specific laboratory tests (vitamins, pre-albumin level etc) may not be reliable in jSSc. It was voted unanimously in both the 2017 and 2018 (21/21) consensus meetings to **include the BMI as a single assessment for response regarding gastrointestinal involvement (Table 1).**

Domain 11: Biomarkers

No peripheral blood biomarker has been fully validated to the extent that it can be used to measure response in a jSSc trial. It was unanimously agreed (21/21) that it is **appropriate to collect biosamples, when possible and available, though a particular serological biomarker(s) was not targeted (Table 1)**.

Domain 12: Growth and Development

In growing children, normal growth and development is important. In the 2017 and 2018 consensus meetings both **delay in sexual maturation and decrease in growth velocity** were considered as potential outcome measures for a jSSc study; however, both were **voted unanimously against being included as outcome measures (Table 1)**. It was felt that there are too many factors that contribute to growth and development (eg. gender, age, nutrition) to be reliable as measures of response to treatment in a jSSc trial.

Discussion

In JIA, guidance for measurements and clinical trials are available(52). The present effort is the first such guidance in jSSc (**Table 1; Supplemental CRFs**). We specifically aimed this proposal at a 12-month, open label jSSc clinical treatment trial. It was not aimed at clinical practice or other trial designs (e.g., double-blind design), because this design is common in pediatric rheumatology and is a simple design to carry through.

This proposal has some significant strengths. It called together diverse medical specialties concerned with SSc as well as patients, and it built on knowledge of the literature (mostly adult SSc and JIA studies). Also, it was developed over several years and included updated data from 3 jSSc registries (Inceptions, CARRA, NPCOS) as well as review of the literature, thus supplying as much factual background as possible and on an ongoing basis. Of note, all the core variables for the composite validated adult SSc outcome measure, CRISS, were captured in our jSSc international consensus (5). This includes 3 out of the 4 "non-response" criteria of Step 1 (only decrease FVC $\geq$ 15% was not included in jSSc, **Table 1**), and all four components of Step 2 (mRSS, FVC%, MD-global, Patient-global). One of the next steps of this group is considering the validation of CRISS in jSSc.

There are also limitations. This was oriented towards a 12-month open label clinical trial and additional considerations would be needed if one were to consider a single blind or double -blind study design. Some

measures were dependent on expert opinion alone (eg. mRSS) and will need validation. Some novel tools in jSSc, such as capillaroscopy and sonography have only been used in jSSc in observational setting and are matter of future research.

The goal in the near future is to pilot these outcomes (**Supplemental CRFs**) in the jSSc cohorts, with particular focus on new or established patients starting medications to evaluate the change of these outcomes in jSSc. Both individual outcomes will be evaluated as well as a composite measure, with options to weigh measures.

Acknowledgments:

Nicola Helmus, Kim Fliglstone, Muthana Al-Obaidi, Michael Blakley, Francesca Ingegnoli, Dana Nemcova, Richard Silver, Alison Zheng, Catherine Orteu, Eileen Baildem, Tamas Constantin, Megan Curran, Antonia Höger, Valentina Leone, Amanda Saracino, Lusine Ambartsumyan, Christian Beyer, Brandi Stevens

References

1. Foeldvari I, Klotsche J, Kasapcopur O, Adrovic A, Terreri MT, Sakamoto AP, et al. Differences Sustained Between Diffuse and Limited Forms of Juvenile Systemic Sclerosis in an Expanded International Cohort. Arthritis Care Res (Hoboken). 2022;74(10):1575-84.
2. Foeldvari I, Torok KS. Review for best practice in clinical rheumatology juvenile systemic sclerosis - Updates and practice points. Best Pract Res Clin Rheumatol. 2021;35(3):101688.
3. Medsger TA, Jr., Silman AJ, Steen VD, Black CM, Akesson A, Bacon PA, et al. A disease severity scale for systemic sclerosis: development and testing. J Rheumatol. 1999;26(10):2159-67.
4. La Torre F, Martini G, Russo R, Katsicas MM, Corona F, Calcagno G, et al. A preliminary disease severity score for juvenile systemic sclerosis. Arthritis Rheum. 2012;64(12):4143-50.
5. Khanna D, Berrocal VJ, Giannini EH, Seibold JR, Merkel PA, Mayes MD, et al. The American College of Rheumatology Provisional Composite Response Index for Clinical Trials in Early Diffuse Cutaneous Systemic Sclerosis. Arthritis Care Res (Hoboken). 2016;68(2):167-78.
6. Valentini G, Bencivelli W, Bombardieri S, D'Angelo S, Della Rossa A, Silman AJ, et al. European Scleroderma Study Group to define disease activity criteria for systemic sclerosis. III. Assessment of the construct validity of the preliminary activity criteria. Ann Rheum Dis. 2003;62(9):901-3.
7. Valentini G, Iudici M, Walker UA, Jaeger VK, Baron M, Carreira P, et al. The European Scleroderma Trials and Research group (EUSTAR) task force for the development of revised activity criteria for systemic sclerosis: derivation and validation of a preliminarily revised EUSTAR activity index. Ann Rheum Dis. 2017;76(1):270-6.
8. Klotsche J, Foeldvari I, Kasapcopur O, Adrovic A, Torok KS, et al. HOW THE ADULT CRISS WORKS IN PEDIATRIC JSSC PATIENTS - RESULTS FROM THE JUVENILE SCLERODERMA INCEPTION COHORT. Annals of Rheumatic Diseases. 2020.
9. Foeldvari I, Klotsche J, Kasapcopur O, Adrovic A, Terreri MT, Sakamoto AP, et al. Differences sustained between diffuse and limited forms of juvenile systemic sclerosis in expanded international cohort. www.juvenile-scleroderma.com. Arthritis Care Res (Hoboken). 2021.
10. Stevens BE, Torok KS, Li SC, Hershey N, Curran M, Higgins GC, et al. Clinical Characteristics and Factors Associated With Disability and Impaired Quality of Life in Children With Juvenile Systemic Sclerosis: Results From the Childhood Arthritis and Rheumatology Research Alliance Legacy Registry. Arthritis Care Res (Hoboken). 2018;70(12):1806-13.
11. Ramey DR, Raynauld JP, Fries JF. The health assessment questionnaire 1992: status and review. Arthritis Care Res. 1992;5(3):119-29.
12. Singh G, Athreya BH, Fries JF, Goldsmith DP. Measurement of health status in children with juvenile rheumatoid arthritis. Arthritis Rheum. 1994;37(12):1761-9.
13. Steen VD, Medsger TA, Jr. The value of the Health Assessment Questionnaire and special patient-generated scales to demonstrate change in systemic sclerosis patients over time. Arthritis Rheum. 1997;40(11):1984-91.
14. Johnson SR, Hawker GA, Davis AM. The health assessment questionnaire disability index and scleroderma health assessment questionnaire in scleroderma trials: an evaluation of their measurement properties. Arthritis Rheum. 2005;53(2):256-62.
15. Ruperto N, Ravelli A, Pistorio A, Malattia C, Cavuto S, Gado-West L, et al. Cross-cultural adaptation and psychometric evaluation of the Childhood Health Assessment Questionnaire (CHAQ) and the Child Health Questionnaire (CHQ) in 32 countries. Review of the general methodology. Clin Exp Rheumatol. 2001;19(4 Suppl 23):S1-9.
16. Poole JL, Williams CA, Bloch DA, Hollak B, Spitz P. Concurrent validity of the Health Assessment Questionnaire Disability Index in Scleroderma. Arthritis Care Res. 1995;8(3):189-93.
17. Brunner HI, Klein-Gitelman MS, Miller MJ, Barron A, Baldwin N, Trombley M, et al. Minimal clinically important differences of the childhood health assessment questionnaire. J Rheumatol. 2005;32(1):150-61.

18. Varni JW, Burwinkle TM, Seid M. The PedsQL 4.0 as a school population health measure: feasibility, reliability, and validity. Qual Life Res. 2006;15(2):203-15.
19. Varni JW, Seid M, Smith Knight T, Burwinkle T, Brown J, Szer IS. The PedsQL in pediatric rheumatology: reliability, validity, and responsiveness of the Pediatric Quality of Life Inventory Generic Core Scales and Rheumatology Module. Arthritis Rheum. 2002;46(3):714-25.
20. Varni JW, Sherman SA, Burwinkle TM, Dickinson PE, Dixon P. The PedsQL Family Impact Module: preliminary reliability and validity. Health Qual Life Outcomes. 2004;2:55.
21. Landgraf JM, Maunsell E, Speechley KN, Bullinger M, Campbell S, Abetz L, et al. Canadian-French, German and UK versions of the Child Health Questionnaire: methodology and preliminary item scaling results. Qual Life Res. 1998;7(5):433-45.
22 Stevens KJ. Working with children to develop dimensions for a preference-based, generic, pediatric, health-related quality-of-life measure. Qual Health Res. 2010;20(3):340-51.
23. Clements P, Lachenbruch P, Siebold J, White B, Weiner S, Martin R, et al. Inter and intraobserver variability of total skin thickness score (modified Rodnan TSS) in systemic sclerosis. J Rheumatol. 1995;22(7):1281-5.
24 Scalapino K, Arkachaisri T, Lucas M, Fertig N, Helfrich DJ, Londino AV, Jr., et al. Childhood onset systemic sclerosis: classification, clinical and serologic features, and survival in comparison with adult onset disease. J Rheumatol. 2006;33(5):1004-13.
25. Foeldvari I, Klotsche J, Kasapcopur O, Terreri MT, Avcin T, Cimaz R, et al. After 24 Months Observation Period the Patients Related Outcomes Improve Significantly in the Juvenile Scleroderma Inceptions Cohorte [abstract. Arthritis Rheum. 2018;70(Suppl 10).
26. Pauling JD, Saketkoo LA, Matucci-Cerinic M, Ingegnoli F, Khanna D. The patient experience of Raynaud's phenomenon in systemic sclerosis. Rheumatology (Oxford). 2019;58(1):18-26.
27. Frantz C, Avouac J, Distler O, Amrouche F, Godard D, Kennedy AT, et al. Impaired quality of life in systemic sclerosis and patient perception of the disease: A large international survey. Semin Arthritis Rheum. 2016;46(1):115-23.
28. Foeldvari I, Klotsche J, Torok KS, Kasapcopur O, Adrovic A, Stanevica V, et al. CHARACTERISTICS OF THE FIRST 80 PATIENTS AT TIMEPOINT OF FIRST ASSESSMENT INCLUDED IN THE JUVENILE SYSTEMIC SCLEROSIS INCEPTION COHORT. WWW.JUVENILESCLERODERMA.COM. Journal of Scleroderma and Related Disorders. 2018;4(1-13).
29. Gladue H, Maranian P, Paulus HE, Khanna D. Evaluation of test characteristics for outcome measures used in Raynaud's phenomenon clinical trials. Arthritis Care Res (Hoboken). 2013;65(4):630-6.
30. Hughes M, Pauling JD. Exploring the patient experience of digital ulcers in systemic sclerosis. Semin Arthritis Rheum. 2019;48(5):888-94.
31 Bruni C, Ngcozana T, Braschi F, Pucci T, Piemonte G, Benelli L, et al. Preliminary Validation of the Digital Ulcer Clinical Assessment Score in Systemic Sclerosis. J Rheumatol. 2019;46(6):603-8.
32 Blagojevic J, Bellando-Randone S, Abignano G, Avouac J, Cometi L, Czirjak L, et al. Classification, categorization and essential items for digital ulcer evaluation in systemic sclerosis: a DeSScipher/European Scleroderma Trials and Research group (EUSTAR) survey. Arthritis Res Ther. 2019;21(1):35.
33. Low AH, Lax M, Johnson SR, Lee P. Magnetic resonance imaging of the hand in systemic sclerosis. J Rheumatol. 2009;36(5):961-4.
34. Martini G, Vittadello F, Kasapcopur O, Magni Manzoni S, Corona F, Duarte-Salazar C, et al. Factors affecting survival in juvenile systemic sclerosis. Rheumatology (Oxford). 2009;48(2):119-22.
35. Bruni C, Buch MH, Furst DE, De Luca G, Djokovic A, Dumitru RB, et al. Primary systemic sclerosis heart involvement: A systematic literature review and preliminary data-driven, consensus-based WSF/HFA definition. J Scleroderma Relat Disord. 2022;7(1):24-32.
36. Cottin V, Brown KK. Interstitial lung disease associated with systemic sclerosis (SSc-ILD). Respir Res. 2019;20(1):13.
37. Volkmann E. Natural history of systemic sclerosis– related interstitial lung disease: How to identify a progressive fibrosing phenotype. Journal of Scleroderma and Related Disorders. 2020;5:31-40.

38. Showalter K, Hoffmann A, Rouleau G, Aaby D, Lee J, Richardson C, et al. Performance of Forced Vital Capacity and Lung Diffusion Cutpoints for Associated Radiographic Interstitial Lung Disease in Systemic Sclerosis. J Rheumatol. 2018;45(11):1572-6.
39. Quanjer PH, Stanojevic S, Cole TJ, Baur X, Hall GL, Culver BH, et al. Multi-ethnic reference values for spirometry for the 3-95-yr age range: the global lung function 2012 equations. Eur Respir J. 2012;40(6):1324-43.
40. Beydon N, Davis SD, Lombardi E, Allen JL, Arets HG, Aurora P, et al. An official American Thoracic Society/European Respiratory Society statement: pulmonary function testing in preschool children. Am J Respir Crit Care Med. 2007;175(12):1304-45.
41. Silver KC, Silver RM. Management of Systemic-Sclerosis-Associated Interstitial Lung Disease. Rheum Dis Clin North Am. 2015;41(3):439-57.
42. Pugnet G, Marjanovic Z, Deligny C, Boussardon K, Benzidia I, Puyade M, et al. Reproducibility and Utility of the 6-minute Walk Test in Systemic Sclerosis. J Rheumatol. 2018;45(9):1273-80.
43. Vandecasteele E, Melsens K, De Keyser F, De Pauw M, Deschepper E, Decuman S, et al. A prospective, longitudinal study evaluating the baseline six-minute walk test as an individual reference value in systemic sclerosis patients. Clin Exp Rheumatol. 2018;36 Suppl 113(4):95-101.
44. Lammers AE, Hislop AA, Flynn Y, Haworth SG. The 6-minute walk test: normal values for children of 4-11 years of age. Arch Dis Child. 2008;93(6):464-8.
45. Miglioretti DL, Johnson E, Williams A, Greenlee RT, Weinmann S, Solberg LI, et al. The use of computed tomography in pediatrics and the associated radiation exposure and estimated cancer risk. JAMA Pediatr. 2013;167(8):700-7.
46. Flynn JT, Kaelber DC, Baker-Smith CM, Blowey D, Carroll AE, Daniels SR, et al. Clinical Practice Guideline for Screening and Management of High Blood Pressure in Children and Adolescents. Pediatrics. 2017;140(3).
47. Khwaja A. KDIGO clinical practice guidelines for acute kidney injury. Nephron Clin Pract. 2012;120(4):c179-84.
48. Martini G, Foeldvari I, Russo R, Cuttica R, Eberhard A, Ravelli A, et al. Systemic sclerosis in childhood: clinical and immunologic features of 153 patients in an international database. Arthritis Rheum. 2006;54(12):3971-8.
49. Gyger G, Baron M. Systemic Sclerosis: Gastrointestinal Disease and Its Management. Rheum Dis Clin North Am. 2015;41(3):459-73.
50. Tu J, Cheung WW, Mak RH. Inflammation and nutrition in children with chronic kidney disease. World J Nephrol. 2016;5(3):274-82.
51. Duggan C, Watkins J, Walker A, . Clinical Assessment of Nutritional Status. In: Decker BC, editor. Nutrition in Pediatrics: Basic Science, Clinical Applications 2008. p. 5-13.
52. Lovell DJ, Ruperto N, Giannini EH, Martini A. Advances from clinical trials in juvenile idiopathic arthritis. Nat Rev Rheumatol. 2013;9(9):557-63.

Figure legend:

Figure 1. Outcomes important for juvenile systemic sclerosis {jSSc) were determined from 2014- 2016, later defined in context to responsiveness in 2017, and ultimately refined to those appropriate for a 12 month clinical trial in jSSc. The final list includes 22 items within 12 domains through voting at in-person consensus meetings.

*2014 respondents: all participants of the paediatric rheumatology email board, the members of the PRES juvenile scleroderma working group and the active participants of the juvenile scleroderma inception cohort project were invited to participate. 70% of the respondents were experienced paediatric rheumatologists (more than 10 years of experience in the field). The mean number of patients followed-up by respondents was 12.3 jSSc patients. Total number of patients follow-up by all respondents is 574. 95% respondents work at academic medical hospitals.

ǂModerated by DF

**Items were also considered in context to the adult Composite Response Index in Systemic Sclerosis (CRISS) was developed by Dinesh Khanna : The American College of Rheumatology Provisional Composite Response Index for Clinical Trials in Early Diffuse Cutaneous Systemic Sclerosis. Arthritis Rheumatol. 2016 Feb;68{2}:299-311. doi: 10.1002/art.39501. PMID: 26808827; PMCID: PMC4826472.

Table 1. Domains and Items suggested as outcome measures for a 12-month clinical trial in Juvenile Systemic Sclerosis (jSSc) from the 2018 International Consensus Meeting.

Physician measured outcomes		
	Metric, range	**Considerations**
GLOBAL DISEASE ACTIVITY		
Physician Global Assessment of Disease Activity (PGA-A)	Visual analog scale 0 – 10 or 0 – 100	• Should take into account past 7 days • Allowable to consider patients features/conditions compared to prior visit • Same physician to assess at study visit for clinical trial
SKIN		
Modified Rodnan Skin Score (mRSS)	Whole number score 0 – 51	• Physical examination at date of study visit • Consider other cutaneous findings in context to scoring children's skin
DIGITAL ULCERS		
Digital Ulcer Clinical Assessment Score (DUCAS)	Number scale 0.5 digit 0 – 19.5	• Physical examination at date of study visit
MUSCULOSKELETAL		
Total active joint count	Whole number score 0 – 75	• Physical examination at date study visit • number of joints that have *either* joint swelling or LOM with pain/tenderness that is considered secondary to jSSc
CARDIAC		
Left ventricular ejection fraction	Echocardiogram value % (30 – 80)	• Echocardiogram closest to date study visit
New onset LV failure	Echocardiogram evaluation Yes/No	• Echocardiogram closest to date study visit
New onset clinically important arrythmia	EKG evaluation Yes/No	• EKG closest to study date
Development of pulmonary arterial hypertension	Echocardiogram evaluation Yes/No	• Echocardiogram closest to date study visit
PULMONARY		
Forced Vital Capacity (FVC)	Pulmonary function test (PFT) value % of predicted (20 – 100)	• PFT closest to study date • Several demographic variables collected to calculate international standard
Diffusion Capacity of the lungs for Carbon monoxide (DLCO)	PFT value % of predicted (20 – 100)	• PFT closest to study date (age eligible) • Hemoglobin collected to determine Hgb-corrected DLCO value
6 Minute Walk test (6MWT)	Walking test with respiratory therapist Meters (0 – 700)	• 6MWT closest to study visit • Lowest SpO2 during the test also important to evaluate desaturation • Forehead or ear probe preferred over finger probe (Raynaud's)
RENAL		
Development of new scleroderma renal crisis (SRC)	Clinical phenotype – present Yes/No	• Blood value abnormalities in setting new hypertension
GASTROINTESTINAL		
Body Mass Index (BMI)	Measurement for pediatrics using Z-scores	• Weight and height used to calculate

	z ≤ -2 is flagged as malnutrition	
Patient Reported Outcomes		
GLOBAL DISEASE ACTIVITY		
Patient Global Assessment of Disease Activity (PtGA-A)	Visual analog scale 0 – 10 or 0 – 100	• Should take into account past 7 days • Parent of child to fill, depending on age (typically 8 years old or greater can self-report) • Must be consistent person scoring over the length of the trial
GLOBAL HEALTH AND FUNCTION		
Childhood Health Assessment Questionnaire (C-HAQ)	Score 0 – 3 (without any difficulty to unable to do) Total score, which is divided among the 8 domains scores which are modified if aids or devices are used	• Patients <16 years old • If child < 8 years old a parent will fill in this form, for 8 years and older, if developmentally appropriate, the child will fill this form • Timeframe - In the past 7 days
Health Assessment Questionnaire (HAQ)	Score 0 – 3 Same scoring system as C-HAQ	• Patients 16 years or older • Traditional HAQ, which has been widely validated
ORGAN SYSTEMS AND GENERAL		
Visual Analog Scales captured in the C-SHAQ and SHAQ	Visual analog scale 0 – 100	• Same questions C-SHAQ and SHAQ, since Childhood version adapted from adult • Patients 16 or older fill out SHAQ • Timeframe - In the past 7 days
Affected by Pain because of Scleroderma	Visual analog scale 0 – 100	• General, global health
Intestinal problems interfered with daily activities	Visual analog scale 0 – 100	• Gastrointestinal domain
Breathing problems interfered	Visual analog scale 0 – 100	• Pulmonary domain
Raynaud's interfered	Visual analog scale 0 – 100	• Raynaud's domain
Finger ulcers interfered	Visual analog scale 0 – 100	• Digital ulcers domain
All the ways (pain, discomfort, limitations daily life, body changes)	Visual analog scale 0 – 100	• General, global health

2014

Open-ended survey important outcomes in jSSc — Domains: 21

52 respondents* — **Items: 131**

SurveyMonkey

2016

Part 1: (56 respondents)

Ranking/significance of 131 from 2014 survey (rank 0-9)

Domains decr. To 13

Items decr. To 49

Part 2: (17 participants)

Pre-meeting Delphi

Rank 49 items (1-9)

Items with median rank 0-4 removed

Part 3: (14 voting participants)‡

Hamburg Meeting/ consensus workshop

Nominal group technique

Items median rank 7-9 included, those ranked 4-7 discussed

Remaining:

Domains: 12

Items: 41

2017

Hamburg Meeting/ consensus workshop

21 voting participants‡

Nominal Group Technique

Voting on the Responsiveness of 41 items from the 2016 Conference**

Remaining:

Domains: 12

Items: 22

2018

Hamburg Meeting/ consensus workshop

22 voting participants‡

Nominal Group Technique

Goal: Consensus for specific items for a 12 month clinical trial in jSSc

Remaining:

Domains: 12

Items: 22