EUROPEAN MEDICINES AGENCY
SCIENCE MEDICINES HEALTH

07 September 2026
EMA/194597/2026
Committee for Medicinal Products for Human Use (CHMP)

# Reflection paper on the clinical investigation of medicinal products for the treatment of systemic sclerosis

Draft

| | |
|---|---|
| Draft agreed by Immune and Inflammatory Diseases Working Party | June 2026 |
| Adopted by PDCO for release for consultation | 24 July 2026 |
| Adopted by CHMP for release for consultation | 7 September 2026 |
| Start of public consultation | 18 September 2026 |
| End of consultation (deadline for comments) | 31 March 2027 |

Comments should be provided using this EUSurvey form. For any technical issues, please contact the EUSurvey Support.


| | |
|---|---|
| Keywords | systemic sclerosis, clinical development |



**Official address** Domenico Scarlattilaan 6 ● 1083 HS Amsterdam ● The Netherlands
**Address for visits and deliveries** Refer to www.ema.europa.eu/how-to-find-us
**Send us a question** Go to www.ema.europa.eu/contact **Telephone** +31 (0)88 781 6000 An agency of the European Union

# Reflection paper on the clinical investigation of medicinal products for the treatment of systemic sclerosis

## Table of contents

# 1. Introduction (background)

*Epidemiology and aetiology*

Systemic sclerosis (SSc), also known as scleroderma, is a rare, chronic, progressive, connective tissue disorder characterised by, among others, autoimmunity, vasculopathy and tissue fibrosis with possible multiorgan involvement. Overall, SSc remains the rheumatic disease with the highest morbidity and mortality, despite recent improvement in survival. In a systematic review and meta-analysis on the prevalence and incidence of SSc across the world, the overall pooled prevalence of SSc was 17.6 (95% confidence interval (CI) 15.1, 20.5) per 100,000 and the pooled incidence rate was 1.4 (95% CI 1.1, 1.9) per 100,000 person-years. In Europe, the pooled prevalence was reported to be 14.8 (95% CI 11.6, 18.8) per 100,000. The pooled incidence and pooled prevalence in women were almost five times higher than in men (Bairkdar et al, 2021).

The pathogenesis of SSc is complex and remains incompletely understood. The dominant hypothesis focuses on the interplay between early immunological events with the production of autoantibodies and vascular changes, which results in the generation of activated myofibroblasts generally considered to be one of the effector cell types in the disease (Lescoat et al, 2025).

*Clinical presentation and disease course*

SSc is a heterogeneous disease, which is reflected in the variety of involved organs and variable disease progression/severity.

Skin involvement (thickening of the skin, scleroderma) is the hallmark feature of SSc. In the initial phases, the fingers of the hands become puffy and over time, excessive collagen deposition leads to skin thickening. This process may spread to other body areas causing limited movement over joints and ultimately contractures of both large and small joints in some patients. In addition to cutaneous fibrosis, approximately 25% of patients with SSc develop cutaneous calcinosis (Volkmann et al, 2023). Next to skin involvement, vasculopathy is almost always present and manifests primarily as Raynaud's phenomenon, which can predate other symptoms of the disease by years. Many patients develop progressive structural changes in the small blood vessels, visible by nail fold capillaroscopy, and around 50% develop digital ulcerations at some time point during the disease (Lescoat et al, 2025).

Vasculopathy is also seen in the lungs, where it can result in pulmonary arterial hypertension (PAH) seen in 10-15% of the patients. PAH is characterized by a progressive increase of pulmonary vascular resistance (PVR) leading to right ventricular failure and premature death. Pulmonary involvement can also manifest as interstitial lung disease (ILD), which is seen in up to 70% of the patients with SSc, and is currently the leading cause of death (Lescoat et al, 2025).

Nearly 90 % of patients with SSc have evidence of gastrointestinal involvement, the oesophagus is the most frequently affected location, but any part of the gastrointestinal system may be involved (Lescoat et al, 2025).

The renal manifestations of SSc are heterogeneous, ranging from mild abnormalities such as microalbuminuria to indirect consequences, including hypertension, in more advanced cases. Renal crisis is a life-threatening condition with acute malignant hypertension, oliguric kidney failure, abnormal urinalysis, and thrombocytopaenia with microangiopathic haemolytic anaemia (Scheen et al, 2023).

Musculoskeletal involvement is characterised by a variety of symptoms, ranging from arthralgia, arthritis, myositis, tendinitis, and contractures to tendon friction rubs; the latter being a hallmark of progressive disease.

Other involved systems include the genitourinary system.

*Standard Treatment*

The management of patients with SSc includes both non-pharmacological and pharmacological interventions and aims at preventing further vascular and/or fibrotic complications of the disease. Due to the wide spectrum of disease manifestations and organ involvement, the management of the disease differs across individual patients. In general, the European alliance of associations for rheumatology (EULAR) recommendations for the treatment of SSc manifestations and updates thereof should be considered (Del Galdo et al, 2025).

# 2. Scope

As a reflection paper, this document provides a high-level description of the requirements for the clinical investigation of medicinal products for the treatment of both specific features of SSc as well as products aimed at a general SSc indication. The regulatory experience with the approval of new medicinal products is limited to SSc-associated disease manifestations. More specific guidance may be developed upon increasing regulatory experience.

# 3. Legal basis and relevant guidelines

This document should be read in conjunction with the introduction and general principles of Annex I to Directive 2001/83/EC, as amended, and all other relevant EU and International Council for Harmonisation (ICH) guidelines (in their current version). These include, but are not limited to:

- Guideline on the clinical investigations of medicinal products for the treatment of pulmonary arterial hypertension (EMEA/CHMP/EWP/356954/2008)
- Guideline on clinical trials in small populations (CHMP/EWP/83561/2005)
- Points to consider on application with 1. meta-analyses; 2. one pivotal study (CPMP/EWP/2330/99)
- ICH E1A Population exposure: the extent of population exposure to assess clinical safety (CPMP/ICH/375/95)
- ICH E7 Note for guidance on studies in support of special populations: geriatrics (CPMP/ICH/379/95)
- ICH M12 Guideline on drug interaction studies (EMA/CHMP/ICH/652460/2022)
- ICH E9 Note for guidance on statistical principles for clinical trials (CPMP/ICH/363/96)
- ICH E9 (R1) Addendum on estimands and sensitivity analysis in clinical trials to the guideline on statistical principles for clinical trials (EMA/CHMP/ICH/436221/2017)
- ICH E17 Guideline on general principles for planning and design of multi-regional clinical trials (EMA/CHMP/ICH/453276/2016 Rev.1)
- ICH M15 Guideline on General principles for model-informed drug development (EMA/CHMP/ICH/496426/2024)
- ICH E11A Guideline on paediatric extrapolation (EMA/CHMP/ICH/205218/2022)
- Reflection paper on the use of extrapolation in the development of medicines for paediatrics (EMA/189724/2018)
- Guideline on missing data in confirmatory clinical trials (EMA/CPMP/EWP/1776/99 Rev. 1)
- Guideline on adjustment for baseline covariates in clinical trials (EMA/CHMP/295050/2013)

- ICH E10 Note for guidance on choice of control group in clinical trials (CPMP/ICH/364/96)
- ICH E4 Dose-Response information to support drug registration (CPMP/ICH/378/95)

# 4. Patient selection

*General considerations for patient selection for SSc studies*

In general, patients selected for inclusion in clinical SSc studies, should be classified according to internationally established criteria such as the American College of Rheumatology (ACR) and EULAR Classification 2013 criteria for SSc (van den Hoogen et al, 2013). It is also expected that the included population is representative for the intended target indication or at least allows generalization to the target population. If the objective of the study is to detect and treat early disease, other inclusion criteria for patients not covered by the ACR/EULAR 2013 criteria may be used (LeRoy et al, 2001; Avouac et al, 2011).

Based on the extent of skin involvement, patients with SSc are clinically classified into two main subtypes: the limited cutaneous SSc (lcSSc) and the diffuse cutaneous SSc (dcSSc) subtypes, represented in respectively two thirds and one third of the SSc patients' population. Since these subtypes differ in prognosis, disease manifestations, and treatments, the study participants should be characterised based on skin subtypes (LeRoy et al, 1988).

The skin subgroups are widely used and endorsed by the majority of experts but do however have some limitations, especially regarding the correct classification of a new patient in the beginning of the disease course (Johnson et al, 2018). Thus, other types of subsets are currently discussed, for example grouping depending on the type of antinuclear antibodies (ANA) profile. ANA is detected in more than 90% of SSc-patients and may present several years before disease onset. Among them, three predominant and specific antibodies are observed: anti-centromere antibodies (ACA), anti-Scl70, and anti-RNA polymerase III (anti-RNA pol III) antibodies, each with different association to disease manifestation and/or disease progression (Elhai et al, 2022).

A small group of patients has SSc *sine* scleroderma, i.e. features of SSc but without any skin involvement and features of SSc can also be seen in other systemic rheumatic diseases, referred to as SSc overlap syndromes.

However, at the time being, subgrouping the patients with respect of skin involvement is considered clinically important. Due to the difference in prognosis, disease manifestations, antibody profile, and treatments, extrapolation of efficacy and safety data between the two subsets based on skin involvement are not considered appropriate by default. Thus, each subtype should be sufficiently represented in a developmental programme to support a global SSc indication comprising both lcSSc and dcSSc.

Because SSc can manifest as a paraneoplastic manifestation of malignancy, cancer screening according to internationally accepted guidelines should be performed prior to study participation (Shah et al, 2015).

*Additional consideration for skin involvement*

Rapid progression of skin involvement has been associated with a worse outcome, with respect to internal organ involvement, and overall survival. Regression of skin involvement is however a characteristic feature of the natural history of skin fibrosis in patients with dcSSc but the time to peak skin score varies widely in patients, leading to a highly heterogeneous pattern of regressing, progressing and stable condition even in early disease stages (Dobrota et al, 2016). Thus, several efforts have been made to define specific features indicative for active skin disease with a high risk for

progression. These features could include a short disease duration, amount of skin involvement (defining a lowest and highest limit), signs of tendon friction rubs, signs of increase in skin involvement during a specific time period, and selected inflammatory markers (Dobrota et al 2016, Herrick et al 2018). This is important information when considering different enrichment strategies and selection of stratification factors for clinical trials.

*Additional considerations for lung involvement*:

Evidence of established or progressing ILD should be documented using High Resolution Computer Tomography (HRCT) or other validated imaging techniques (Antoniou et al, 2026). It is recommended that the ILD pattern is determined, as several forms may occur in SSc and treatment responses may vary between them. Further, it is recommended that the extent of ILD is defined, and that other possible reasons for impaired lung-function test i.e. Forced expiratory volume (FEV)/ Diffusing capacity of the lungs for carbon monoxide (DLCO) are excluded.

# 5. Assessment of efficacy

## *5.1. Efficacy criteria/treatment goals*

### 5.1.1. Different types of potential outcome measures and endpoints

For a new medicinal product intended for SSc, treatment goals may target either the overall disease course or specific disease manifestations. A specific disease subtype or both major disease subtypes could be targeted.

Treatment objectives include improvement of organ function and/or symptoms or delay of onset or prevention/slowing of progression of clinical manifestations. Disease modification may be included as a treatment goal, depending on the mechanism of action, if a clear and appropriately justified definition of this term is provided and an appropriate (validated) outcome measure is proposed. How this prerequisite can be fulfilled would need to be thoroughly justified in development programs that includes this treatment goal.

Selection of primary, (key) secondary, and exploratory endpoints should be based on goals of therapy and target indication, taking into consideration the heterogeneity of the disease manifestations. The heterogeneity includes (but is not limited to) the variety of potential internal organs involved. Potential endpoints need to be sufficiently validated and fit for purpose. This generally applies for potential new endpoints, including endpoints based on digital health technologies.

If surrogate variables (for e.g. involvement of internal organs or for survival) are proposed, the strength of evidence for surrogacy would need to be carefully considered and justified, please refer to ICH E9 Note for guidance on statistical principles for clinical trials (CPMP/ICH/363/96). Relationships between clinical and surrogate variables for one product do not necessarily apply to a product with a different mode of action for treating the same disease.

Potential endpoints include:

- Endpoints/measures related to specific disease (organ-) manifestations
- Physicians and patient-reported global assessments and other Patient Reported Outcomes (PROs)
- Dedicated composite endpoints for subpopulations or for a broad overall SSc indication
- Overall survival

Data showing efficacy on a specific clinical manifestation typically associated with SSc generated in studies from non-SSc indications could be supportive of a SSc indication, if fully justified on a case-by-case basis (see also Section 6 of this document). Possible examples for which this may apply is efficacy data on PAH and ILD. Evidence for target indications not specific for the SSc could come from studies conducted in a broader population including an adequately sized subgroup of patients with SSc. However, in the absence of a robust justification, dedicated studies for a SSc population are expected to be performed.

*Skin fibrosis endpoints*

mRSS is an established outcome measure for evaluating skin fibrosis that assesses skin involvement in different anatomical sites. Its relevance in predicting disease outcome has been shown (Del Galdo et al, 2020). Increase in mRSS over time has been associated with poor clinical outcomes in SSc. However, the use of mRSS response as surrogate of organ changes is controversial (Pope et al, 2023).

mRSS is a relevant outcome measure for skin fibrosis also without supportive histologic or imaging data. Reliability depends on investigator's experience and consistency of assessments across raters, and thus, standardised training of investigators and a single investigator to perform the assessment at each scheduled evaluation during the study is recommended as well as periodic central quality review.

Other endpoints than those based on mRSS may be of relevance for evaluating subjects predominantly affected by skin manifestations, see e.g. section on physician and patient-reported assessments below.

*Digital ulcer endpoints*

An endpoint measuring the number of new digital ulcers from baseline is adequate for digital ulcer assessment.

Time to healing of digital ulcers can also be measured. In relation to evaluation of digital ulcers, pain assessment is of critical importance.

*ILD endpoints*

There is a wide range of endpoints that could potentially be used to assess SSc-associated ILD. The measures provided below should not be understood as a complete list but rather as examples.

Forced vital capacity (FVC) and DLCO are appropriate outcome measures for SSc associated ILD (Moore, 2012). Continuous endpoints based on FVC as well as percentage of responders may be of value.

The St George's Respiratory Questionnaire (SGRQ) is a questionnaire that was initially developed in asthma and Chronic Obstructive Pulmonary Disease (COPD) but has also been validated for use in e.g. ILD (City St George's, University of London).

Radiological findings i.e. changes in radiological structures as measured with HRCT are also of value to evaluate as descriptive measures.

*PAH endpoints*

6-minute walk distance (reflecting exercise capacity) and haemodynamic measurement can be used to assess SSc-associated PAH. Composite endpoints that reflect time to clinical worsening can be used as well.

In line with the Guideline on the clinical investigations of medicinal products for the treatment of pulmonary arterial hypertension (EMEA/CHMP/EWP/356954/2008), the place of haemodynamic measurements is currently limited to the diagnosis and as a secondary endpoint. Haemodynamic measures may play an important role during the early development of the medicinal product to

elucidate the mechanism of action or to define the dose-response relationship. The role of echocardiography and other imaging techniques in estimating haemodynamic variables may offer potential non-invasive tools to evaluate the efficacy of the medicinal intervention.

Another possible outcome measure is World Health Organisation (WHO) functional class improvement.

*Endpoints related to other disease manifestations*

New incidence, worsening or improvement of other key internal organ involvement should systematically be captured, such as renal (renal crisis), liver, cardiac/cardiovascular, and gastrointestinal (GI). Clinical assessments on heart and renal involvement are recommended, in order to ensure that potential effects of a new medicinal product on the key features of the disease are measured and quantified.

Assessment of musculoskeletal manifestations (including arthritis, myositis, tendon friction rubs, contracture progression prevention) is valuable.

To assess GI involvement, the GI involvement scale (University of California Los Angeles Scleroderma Clinical Trial Consortium Gastrointestinal Tract (UCLA SCTC GIT)) has been developed. This is a self-reported GI quality of life 7–multi-item scale specifically intended for the range of problems that can occur in SSc with areas of reflux, distention/bloating, diarrhoea, faecal soilage, constipation, emotional well-being, and social functioning (Pope, 2011).

For Raynaud phenomenon, the Raynaud Condition Score (RCS) is available. The RCS is a self-reported global assessment of Raynaud's phenomenon activity using a 0 –10 ordinal scale, which incorporates the cumulative daily frequency, duration, severity, and impact of RP attacks. A composite score from daily measures is then calculated (Pope, 2011).

*Physicians and patient-reported global assessments and other PROs*

Outcomes based on physician-reported as well as patient-reported global assessment of overall health may be relevant.

Regarding overall impact on health status, the use of both disease manifestation specific and general assessments of quality of life may prove useful. For comments on PROs related to specific organ manifestations of SSc, see above headings.

The Health Assessment Questionnaire–Disability Index (HAQ-DI) and Scleroderma Health Assessment Questionnaire (SHAQ) are possible PROs for SSc studies. The HAQ-DI is a self-reported questionnaire in 8 domains, not specific for SSc, that measures self-reported functions. The SHAQ consists of the 8 HAQ domains and also includes the following scales: pain, patient global assessment, vascular, digital ulcers, lung involvement, and GI involvement (Pope, 2011).

For assessment of self-reported fatigue and its impact upon daily activities and function, the Functional Assessment of Chronic Illness Therapy – Fatigue Scale (FACIT-Fatigue) can be used (FACIT group).

Pain-assessment should be based on Visual Analogue Scale (VAS)- or Numeric Rating Scale (NRS)-scales and is a relevant PRO especially for skin manifestations, e.g. digital ulcers.

Also assessing PROs, in which productivity at work and social integration are targeted could be considered.

Finally, also other PROs, including newly developed PROs may be considered, if sufficiently validated. A general approach may be to validate new outcome measures using data from earlier phases of the development programme (e.g. phase 2) to be able to use the outcome measure to collect confirmatory evidence in later phases (phase 3).

*Composite outcome measures combining multiple disease domains*

Several composite outcome measures combining multiple disease domains have been developed (Del Galdo et al, 2020) that may be used as outcome measure if sufficiently validated (see Section 5.1.2).

The provisional American College of Rheumatology Composite Response Index in Systemic Sclerosis (ACR-CRISS) is a composite endpoint that captures the global evaluation of the likelihood of improvement in early SSc. It is derived from a two-step algorithm, with probabilities of improvement based on five core set measures: mRSS, FVC % predicted, HAQ-DI, the Patient Global Assessment (PGA), and the Clinician Global Assessment (CGA)—each of which are weighted differently in each study. To address limitations of the ACR-CRISS, the revised CRISS (rCRISS) was developed (Khanna et al, 2021).

*Assessment of effects on survival*

Effects on overall survival should be captured and presented in the marketing authorisation application of a new medicinal product for SSc.

For novel treatments for SSc, assessment of the impact on mortality, is important. It is acknowledged that to reliably assess this, randomised clinical studies of longer duration with a larger number of patients are needed. Nonetheless, even if a claim of prolonged survival is not made, effects on mortality should be captured and reported at least in a descriptive manner. This can be done in an open-label extension phase, to supplement data collection in a controlled part of a trial.

### 5.1.2. Choice of primary endpoint and secondary endpoints for different target indications

For an **indication targeting SSc-ILD**, a primary endpoint that includes a FVC-measurement is appropriate. Change from baseline in percentage FVC predicted measured over a sufficiently long time period, at least 52 weeks, is recommended. Percentage predicted FVC standardises the absolute measurement and reduces or eliminates variability due to demographic factors such as age, gender, body size which could be different between treatment arms in a clinical study.

The proportion of subjects with a specified absolute FVC decline in percent predicted e.g. 5% is relevant to present among the secondary endpoints as well as FVC expressed as absolute change from baseline. DLCO-measurements are also recommended to be included among the secondary endpoints, preferably among endpoints included in a testing procedure that controls the type I error. There are also other potential secondary endpoints, please refer to Section 5.1.1 above.

For an **indication targeting PAH** secondary to scleroderma, refer to the Guideline on the clinical investigations of medicinal products for the treatment of pulmonary arterial hypertension (EMEA/CHMP/EWP/356954/2008).

For an **indication targeting digital ulcers** in patients with SSc, a primary endpoint measuring the number of new digital ulcers from baseline to study endpoint is adequate and should be supported by pain assessments. Measuring such an endpoint after 4-6 months of treatment may be reasonable but the timepoint will depend on the mechanism of action of the medicinal product under investigation. Assessment of pain associated with digital ulcers is important to adequately characterise patient burden.

For an **indication targeting skin fibrosis**, a measurement based mRSS is an appropriate primary endpoint. If mRSS is measured as change from baseline, it should be supplemented by responder assessment on mRSS on the basis of a clinically justified response threshold (i.e. either in terms of patient-perceived benefit or predicting improved longer-term/systemic disease), and by other clinically

relevant outcomes that may include assessment of other specific organ involvement. With regards to timing of the mRSS, measuring change from baseline in at week 52 is appropriate. PROs specifically targeting skin disease may be used provided that the outcome measures are adequately validated and fit for purpose.

Regulatory experience with the rCRISS is currently limited. Nevertheless, rCRISS may be considered as a primary outcome measure for an **indication targeting early dcSSc**, provided that sufficient evidence demonstrating its validation for this specific use is available and can be presented. Data supporting the interpretability of the endpoint are critical. A transparent and comprehensive presentation of the results for each individual component of rCRISS is mandatory and these should preferably be included as separate (secondary) endpoints. In the absence of consistent positive effects across the individual components, a thorough and critical evaluation of the overall outcome would be required.

As an alternative to a composite endpoint combining multiple disease domains as the primary endpoint to support a broad, **non-manifestation specific indication, potentially including both dcSSc and lcSSc,** a co-primary endpoint of different outcome measures could be used. In case such strategy for the primary endpoint is selected, interaction with the European Medicines Agency (EMA), through i.e. a CHMP Scientific Advice procedure, is recommended.

### *5.2. Methods to assess efficacy criteria*

The outcome measures mentioned in Section 5.1 should be assessed according to established and validated methods.

The optimal timing of the efficacy outcome assessments depends on e.g. the mode of action of the medicinal product to be investigated. For the timing of the evaluation in the confirmatory study, there should be data available from earlier phases of the development programme to guide the choice. This may also include data on the PK and PD characteristics of the medicinal product.

Some comments on timepoints for outcome assessments are provided in Section 5.1 of this document, under the headings that specifically mention different endpoints and choice of endpoints. For relatively short-term endpoint measurements, some additional support of sustained efficacy over time may still be needed. The nature of this support can vary but the strategy needs to be well justified. Please refer also to Section 6 for guidance on optimal duration of studies with the objective to collect confirmatory evidence.

## 6. Study design

### *6.1. Pharmacology studies*

#### 6.1.1. Pharmacokinetics

The Pharmacokinetic (PK) properties of the medicinal product should be thoroughly investigated in accordance with relevant guidelines.

#### 6.1.2. Pharmacodynamics

The pathogenesis of SSc involves three interrelated processes: vasculopathy, immune dysregulation with the production of autoantibodies targeting nuclear antigens and fibrosis (Lescoat et al, 2025).

Depending on the specific target and the mechanism of action of the medicinal product under development, pharmacodynamic (PD) endpoints will have to be selected accordingly. The relevant PD endpoints should be validated and their role in the pathophysiology of the SSc should be previously established. Early regulatory interaction through central EMA Scientific Advice request is recommended, in particular if the proposed PD endpoint is planned to be used also in the phase 3 study as a mean to collect supportive evidence of effect.

### 6.1.3. Drug-drug interactions (DDI)

Interaction studies should be performed in accordance with current guidelines (see ICH M12 Guideline on drug interaction studies). Evaluation of any interaction potential should be performed with medicinal products likely to be co-administered in clinical practice or planned to be co-administered during clinical studies. In the treatment of SSc, this in particular concerns treatments of different organ manifestations. Each of these manifestations often requires treatment with different classes of drugs. The need for conducting interaction studies should be based on the known PK and PD properties of the medicinal product.

## *6.2. Therapeutic studies*

### 6.2.1. Exploratory and dose finding studies

Following proof of concept, dose finding studies in patients with SSc are recommended to identify the optimal posology regimen for the phase 3 development. The population included in these early studies should inform on the SSc target population of the confirmatory study(ies). A minimum effective dose and dose response curve are preferably established. Efforts should be undertaken to explore different doses or intervals according to the respective patient characteristics (i.e. disease severity and subtype, clinical manifestations, special populations in need of lower doses) as well as to define the need for weight adjustment or adjustment to other co-variates.

The general principles for model-informed drug development should be taken into account, see ICH M15 General principles for model-informed drug development.

For dose finding, placebo-controlled parallel group studies are recommended, if feasible.

### 6.2.2. Confirmatory studies

***Design elements***

Considering the rarity of the condition, in most cases relying on a single pivotal study is possible. This may be acceptable provided that the study is particularly compelling with respect to internal and external validity, clinical relevance and data quality. The heterogeneity of this condition is a challenge in clinical study design.

Generally, a randomised double-blind placebo-controlled parallel group design to demonstrate the superiority of the new treatment over placebo is recommended for generating confirmatory evidence of efficacy in SSc.

If, in some particular cases, placebo-controlled design of the study is considered not ethical or feasible, an alternative design could be considered. For example, a randomised, active-controlled design for novel therapies in this rare disease setting could be potentially acceptable. The blinding will be required as feasible; however, if some open label design elements are considered indispensable, efforts for a blinded efficacy assessment should be made. Regulatory interaction through central EMA

Scientific Advice request is recommended early during the development, in particular if the double-blind placebo-controlled design is not planned.

The selection of stratification factors should be made according to their impact on prognosis or on treatment responsiveness. Randomisation should be stratified by region (ICH E17 Guideline on general principles for planning and design of multi-regional clinical trials) and other important stratification factors relevant for the clinical setting should be considered, e.g. disease duration, background immunosuppressive therapy, disease subsets, organ manifestations, etc. (please see also Section 4 Patient selection).

Different treatments have been developed to treat the specific manifestations of the SSc (e.g. PAH, ILD, etc) due to their selective mechanism of action. For other treatments with broader effect (e.g. anti-CD20, cellular therapy), there may be a need to assess and demonstrate efficacy in different organs/outcomes within the single phase 3 study. For potential treatment goals, please see the Section 5.1 above. The treatment period of the confirmatory study should generally be at least of 52-week duration to allow assessing maintenance of efficacy and collecting double-blind data, with an adequate follow-up extension phase. In particular, in placebo-controlled designs, study subjects need to have access to standard of care and should be able to receive full symptomatic treatment for their various organ manifestations. Further, appropriate criteria for initiating rescue therapy is recommended to be in place.

An early request for EMA Scientific Advice is recommended when adaptive study designs, such as seamless phase II/III designs, are planned. The acceptability of such approaches will be assessed on a case-by-case basis.

***Patient selection/target population***

The patient population should be diagnosed according to the internationally established criteria, please refer to the Section 4 of this document. The target population should be selected according to the treatment goals (see Section 5.1 of this document).

Current cohort enrichment strategies for SSc studies focus on a narrow subgroup of patients with SSc (i.e. early diffuse cutaneous SSc patients and/or specific antibody pattern), potentially hindering patient accrual and limiting the generalisability of the study findings (Volkmann et al, 2025). Although these strategies are widely used, they may possibly only support a rather restrictive future indication. Each subtype should be sufficiently represented in a developmental programme to support a claimed broad indication (please refer to the Section 4 of this document).

Moreover, the SSc disease manifestations have different trajectories (including rate of progression) both within one SSc subtype and when comparing dcSSC with lcSSc. This is important to consider when selecting the patient population based on disease duration.

Evidence for target indications not specific for the SSc (e.g. PAH, ILD) could potentially come from studies conducted in a broader population including an adequately sized subgroup of patients with SSc (frequently including wider connective tissue disorders population). However, in the absence of a robust justification, dedicated studies for a SSc population are expected to be performed as this is the preferred option to collect data in PAH associated with SSc.

Standard of care should be well defined in the study protocol, as well as prohibited and rescue medications. The concomitant standard therapy should be in line with applicable treatment guidelines (e.g. Del Galdo et al, 2025), carefully documented, and its impact on results analysed based on a pre-specified plan. Currently, most therapeutic options are off-label and preferences as well as their line in therapy are predominantly experienced-based rather than based on strong evidence. Patients could

often remain on stable immunosuppressive background therapies during the study. Also, the previous use and response to standard therapy should be documented. For clinical development in special populations with SSc, please see below the Section 8 of this document.

***Statistical considerations and estimands***

The statistical analyses should be predefined and aligned with the clinical questions of interest (estimands). A primary estimand for the study should be carefully defined including identification of all relevant intercurrent events (ICE) as per ICH E9(R1) Addendum on estimands and sensitivity analysis in clinical trials to the guideline on statistical principles for clinical trials. Detailed information should be made available to decide on the most appropriate estimand strategy for each ICE. Although the treatment policy estimand is generally acceptable in superiority trials, its adequacy should be evaluated separately for each ICE. Analyses addressing alternative ways to handle intercurrent events can assist in interpretation of study data and assessing robustness of conclusions, thereby supplementing benefit-risk assessment. If the composite strategy is proposed for a responder endpoint, it is advised that the supplementary analysis based on treatment policy strategy is also conducted. Efforts should be made to collect all relevant data for the primary and important other estimands (e.g. follow-up regardless of ICE). The frequency and timing of the different intercurrent events should be reported separately per treatment arm.

For selection of primary and secondary endpoints and timing of endpoints, please refer to Section 5 of this document.

# 7. Safety aspects

## *7.1. Specific effects*

The full immunomodulatory effects, preferably including downstream effects on the immune system function, of the medicinal product under investigation should be characterised, as applicable. Where applicable for biological medicinal products, assessment of immunogenicity, as well as the potential implications of immunogenicity for safety (as well as efficacy) should be characterised (see also below).

For medicinal products administered via subcutaneous and/or intravenous routes, injection-site reactions and infusion-related reactions should be carefully monitored, also in the context of immunogenicity. This includes the assessment of anti-drug antibodies (ADA) and, where applicable, neutralising antibodies (nAb). The potential relationship between the development of ADA and/or nAb and loss of efficacy, infusion-related reactions, and other adverse events should be systematically evaluated. In addition, factors influencing the development of nAb, such as treatment duration, dose, and concomitant medicinal products, should be systematically analysed.

Adverse events (AEs) related to the mechanism of action of the medicinal product and risks known for the relevant substance class should be systematically investigated.

AEs of special interest commonly include infections, notably severe, opportunistic, and common infections. The relationship between the occurrence of infections and relevant immunological parameters, such as neutrophil counts and immunoglobulin levels, should be analysed. In addition, other potential AE of special interest, including malignancies, should be specifically monitored and reported. Consideration should be given to the implementation of an independent adjudication process for the assessment of selected safety outcomes.

Information on prior and concomitant therapies used in clinical studies should be comprehensively documented to explore any influence on the safety profile of the investigational medicinal product. In

addition, organ manifestations of SSc, relevant comorbidities, and cardiovascular risk factors should be considered in the safety evaluation.

## *7.2. Long-term effects*

SSc is a chronic condition requiring long-term treatment; therefore, sufficiently sized controlled safety data covering a minimum treatment duration of 52 weeks should be provided. In general, the extent of patient exposure should be consistent with the requirements outlined in ICH E1A Population exposure: The Extent of Population Exposure to Assess Clinical Safety.

Medicinal products with immunosuppressive mechanisms of action are frequently developed across multiple rheumatological or systemic autoimmune conditions, and, under certain conditions, safety data may be partially extrapolated between related indications. Such extrapolation would only be acceptable if there is sufficient similarity between the conditions with respect to comorbidities, concomitant medication, and other factors influencing the risk of adverse reactions, and that comparable posology, treatment duration, and therapeutic strategies are applied.

Given the progressive yet serious nature of SSc, and the potential for AEs to emerge or persist after treatment discontinuation, long-term safety follow-up is warranted. Furthermore, the reversibility of AEs following treatment withdrawal should be characterised. Participation in disease-specific registries in the post-authorisation setting is advisable.

## *7.3. Safety endpoints*

Safety considerations in SSc trials are particularly complex and should explicitly address the risks of organ-specific complications (e.g. pulmonary, renal, and cardiac), immunosuppression-related adverse events, and the challenge of distinguishing disease progression from treatment-related toxicity.

Clinical safety observations should be supported by appropriate laboratory assessments relevant to the medicinal product under development. Where the development of antibodies is anticipated (see Section 7.1 of this document), the immunogenicity testing methodologies for ADA and nAb should be appropriately validated, and clinically relevant thresholds for a potential impact should be defined.

# 8. Studies in special populations

## *8.1. Studies in elderly patients*

SSc is a condition that can affect all age categories, but late-onset SSc has some distinct clinical characteristics. Separate studies in the elderly are not required but clinical studies should include representative numbers of patients in the age categories 65 – 74, 75 – 84, and 85 years and above, if possible. Arbitrary upper age cut-offs for inclusion should be avoided. It is referred to ICH E7 Note for Guidance on Studies in Support of Special Populations: Geriatrics.

### ***Efficacy in elderly patients***

Efficacy data should be presented for the different relevant age strata across the geriatric population to enable assessment on whether the effects in these groups are consistent with the effects in the non-geriatric population. Age-related subgroup analyses, if allowed by sample-size, are to be pre-specified.

***Safety in elderly patients***

Especially regarding safety, data derived from a younger population may not fully be generalisable to the geriatric population. The risks for co-morbidities such as cardiovascular disorders and renal impairment increases with age, and concomitant drug therapies and risk of drug interaction require special considerations. Elderly patients may also be more susceptible to infections when treated with immune-modulating medicinal products. The clinical development programme, including PK studies, should provide data that allow an assessment whether special warnings and precautions or dose recommendations would be applicable in elderly patients. Available data should be reported separately for patients aged 65-74, 75-85 and 85 and older.

## *8.2. Studies in paediatric patients*

As in the adult form, juvenile Systemic Sclerosis (jSSc) has different clinical subtypes, i.e. diffuse, limited, *sine* scleroderma, or representing overlap syndrome. The clinical presentation is heterogeneous which is reflected in the variety of involved organs and disease progression/severity. SSc manifesting before the age of 16 (juvenile Systemic Sclerosis; jSSc) is rare with an estimated incidence in epidemiological studies in the European Union of 0.27 - 0.5 cases per million children per year (Herrick et al, 2010).

There have been several initiatives to provide treatment guidelines, including the SHARE recommendations for jSSc (Foeldvari et al, 2021). The EULAR recommendations and treatment algorithm are considered guiding as well (Kowal-Bielecka et al, 2017; Fernández-Codina et al, 2018). A significant unmet medical need exists in the jSSc population.

Classification criteria for jSSc can follow the 2013 EULAR/ACR **classification criteria** for adults (van den Hoogen et al, 2013). Although these criteria have not (yet) formally been validated in children, the 2013 EULAR/ACR classification criteria are reported to have higher sensitivity compared to the PRES/ACR/EULAR classification criteria for jSSc (Zulian et al, 2007). In studies including both adults and paediatric patients, uniformity in classification criteria is encouraged.

Direct comparison of disease characteristics between **juvenile and adult SSc** shows that disease phenotype, internal organ involvement, and autoantibody profiles differ across the juvenile and adult groups. Paediatric patients tend to have a higher prevalence of the diffuse cutaneous subtype, while adults more commonly present with the limited form. Vital organ involvement, considered as the main morbidity and mortality factor, is more common among adults, as are scleroderma-specific autoantibodies (a.o. anti-Scl 70); overall survival is poorer in adults compared to children. Musculoskeletal complaints, i.e. arthralgia and arthritis, and finger ulcers are more common among juvenile patients compared to adults (Foeldvari et al, 2021; Adrovic et al, 2023).

In general, therefore, evaluation of efficacy and safety by age-related subgroups would be the most robust approach to account for differences in disease characteristics and potential differences in treatment response in jSSc. Given the rarity of the condition, however, the feasibility of conducting adequately statistically powered, stand-alone clinical efficacy studies in this vulnerable paediatric population is expected to be limited.

For characterising the pharmacological effects of medicines in children, the use of Model Informed Drug Development (MIDD) approaches (population PK, PK/PD, Drug Exposure-Response) should be envisaged. These methods have substantial potential to efficiently analyse data, to robustly support optimisation of clinical trial design, to adequately inform dose selection and optimisation, extrapolation across populations and indications, ultimately facilitating more efficient and

timely paediatric drug development. The general principles outlined in the ICH M15 Guideline on General principles for model-informed drug development should be applied for this purpose.

In addition, the ICH E11A Guideline on paediatric extrapolation provides guidance on extrapolation between adults and paediatric patients, including the identification of uncertainties in the extrapolation concept and development of an extrapolation plan intended to address them. The paediatric clinical development that would suffice to address critical uncertainties, whilst being operationally feasible, should be determined on a case-by-case basis, taking into account product-availability of specific PK characterisation and PK/PD relationship. Disease heterogeneity and differences between adult and paediatric populations with respect to organ involvement, disease course, prognosis, and standard-of-care treatment strategies may have implications for **endpoint selection** and interpretation particularly in paediatric patients below the age of 12 years. As for adults, selection of endpoints should e.g. reflect the study population (in terms of organ involvement) and the indication the Applicant is targeting to claim. The Juvenile Systemic Sclerosis Severity Score (J4S), which was developed by La Torre et al, 2012 is an endpoint tailored to the paediatric population. It includes indices of nine organ systems; each scored on a scale ranging from 0 to 4 according with the severity of each organ involvement. The J4S has been reported to be effective and reliable in detecting change in disease severity. In the absence of fully validated endpoints in paediatric patients, the use of validated endpoints for adults can be considered (please refer to Section 5.1 of this document). The inclusion of endpoints addressing musculoskeletal manifestations is specifically encouraged given the high frequency of occurrence in paediatric compared to adult patients.

Inclusion of adolescents with jSSc in clinical studies conducted in adults is generally recommended, unless scientifically and ethically justified concerns would preclude to do so. Stratification by age could be used. A separate analysis showing consistent effect is required. The level of evidence supporting a favourable benefit–risk balance in adults that would be required prior to inclusion of younger paediatric patients should be carefully evaluated. A staggered enrolment approach might be considered. The timing of paediatric inclusion in the development programme should be agreed in advance with regulatory authorities.

Gastrointestinal involvement may lead to malnutrition and failure to thrive; this needs to be accounted for in paediatric dose strategies.

# 9. References


Adrovic, A., et al. "Juvenile and adult-onset scleroderma: different clinical phenotypes." *Seminars in arthritis and rheumatism*. Vol. 60. WB Saunders, 2023.

Antoniou, Katerina, et al. "ERS/EULAR clinical practice guidelines for connective tissue disease-associated interstitial lung disease". *European Respiratory Journal* 67.1 (2026): 2402533.

Avouac, Jérôme, et al. "Preliminary criteria for the very early diagnosis of systemic sclerosis: results of a Delphi Consensus Study from EULAR Scleroderma Trials and Research Group." *Annals of the rheumatic diseases* 70.3 (2011): 476-481.

Bairkdar, Majd, et al. "Incidence and prevalence of systemic sclerosis globally: a comprehensive systematic review and meta-analysis." *Rheumatology* 60.7 (2021): 3121-3133.

City St George's, University of London, *St George's Respiratory Questionnaire (SGRQ)* (n.d.), https://www.staff.sgul.ac.uk/about/research-operations/research-administration/st-georges-respiratory-questionnaire. Accessed 22 December 2025.

Del Galdo, Francesco, Collette Hartley, and Yannick Allanore. "Randomised controlled trials in systemic sclerosis: patient selection and endpoints for next generation trials." *The Lancet Rheumatology* 2.3 (2020): e173-e184.

Del Galdo, Francesco, et al. "EULAR recommendations for the treatment of systemic sclerosis: 2023 update." *Annals of the rheumatic diseases* 84.1 (2025): 29-40.

Dobrota, Rucsandra, et al. "Prediction of improvement in skin fibrosis in diffuse cutaneous systemic sclerosis: a EUSTAR analysis." *Annals of the rheumatic diseases* 75.10 (2016): 1743-1748.

Elhai, Muriel, et al. "Stratification in systemic sclerosis according to autoantibody status versus skin involvement: a study of the prospective EUSTAR cohort." *The Lancet Rheumatology* 4.11 (2022): e785-e794.

FACIT group, *Functional Assessment of Chronic Illness Therapy (FACIT) – Fatigue Scale*. https://www.facit.org/measures/facit-fatigue. Accessed 22 December 2025.

Fernández-Codina, Andreu, et al. "Treatment algorithms for systemic sclerosis according to experts." *Arthritis & Rheumatology* 70.11 (2018): 1820-1828.

Foeldvari, Ivan, et al. "Are diffuse and limited juvenile systemic sclerosis different in clinical presentation? Clinical characteristics of a juvenile systemic sclerosis cohort." *Journal of Scleroderma and Related Disorders* 4.1 (2019): 49-61.

Foeldvari, Ivan, et al. "Consensus-based recommendations for the management of juvenile systemic sclerosis." *Rheumatology* 60.4 (2021): 1651-1658.

Greco, Raffaella, et al. "Innovative cellular therapies for autoimmune diseases: expert-based position statement and clinical practice recommendations from the EBMT practice harmonization and guidelines committee." *EClinicalMedicine* 69 (2024).

Herrick, Ariane L., et al. "Incidence of childhood linear scleroderma and systemic sclerosis in the UK and Ireland." *Arthritis Care & Research: Official Journal of the American College of Rheumatology* 62.2 (2010): 213-218.

Herrick, Ariane L., et al. "Patterns and predictors of skin score change in early diffuse systemic sclerosis from the European Scleroderma Observational Study. " *Ann Rheum Dis*. Apr;77(4) (2018):563-570.

Johnson, Sindhu R., et al. "There is a need for new systemic sclerosis subset criteria. A content analytic approach." *Scandinavian journal of rheumatology* 47.1 (2018): 62-70.

Khanna, Dinesh, et al. "New composite endpoint in early diffuse cutaneous systemic sclerosis: revisiting the provisional American College of Rheumatology Composite Response Index in Systemic Sclerosis." *Annals of the rheumatic diseases* 80.5 (2021): 641-650.

Kowal-Bielecka, Otylia, et al. "Update of EULAR recommendations for the treatment of systemic sclerosis." *Annals of the rheumatic diseases* 76.8 (2017): 1327-1339.

La Torre, Francesco, et al. "A preliminary disease severity score for juvenile systemic sclerosis." *Arthritis & Rheumatism* 64.12 (2012): 4143-4150.

LeRoy EC, et al. "Scleroderma (systemic sclerosis): classification, subset and pathogenesis. " *Rheumatol.* Feb;15(2) (1988): 202-5

LeRoy, E. Carwile, and Thomas A. Medsger. "Criteria for the classification of early systemic sclerosis." *The Journal of rheumatology* 28.7 (2001): 1573-1576.

Lescoat, Alain, et al. "Systemic sclerosis: pathogenic mechanisms and their implications for treatment." *Seminars in Immunopathology*. Vol. 47. No. 1. Springer Berlin Heidelberg, 2025.

Moore, V. C. "Spirometry: step by step." *Breathe* 8.3 (2012): 232-240.

Pope, Janet. "Measures of systemic sclerosis (scleroderma): Health Assessment Questionnaire (HAQ) and Scleroderma HAQ (SHAQ), physician-and patient-rated global assessments, Symptom Burden Index (SBI), University of California, Los Angeles, Scleroderma Clinical Trials Consortium Gastrointestinal Scale (UCLA SCTC GIT) 2.0, Baseline Dyspnea Index (BDI) and Transition Dyspnea Index (TDI)(Mahler's Index), Cambridge Pulmonary Hypertension Outcome Review (CAMPHOR), and Raynaud's Condition Score (RCS)." *Arthritis care & research* 63 (2011): S98-111.

Pope, Janet E., et al. "State-of-the-art evidence in the treatment of systemic sclerosis." *Nature reviews rheumatology* 19.4 (2023): 212-226.

Royle JG, Lanyon PC, Grainge MJ, et al. The incidence, prevalence, and survival of systemic sclerosis in the UK Clinical Practice Research Datalink. *Clin Rheumatol.* 2018;37:2103–11.

Scheen, Marc, et al. "Renal involvement in systemic sclerosis." *Autoimmunity reviews* 22.6 (2023): 103330.

Shah, Ami A., and Livia Casciola-Rosen. "Cancer and scleroderma: a paraneoplastic disease with implications for malignancy screening." *Current opinion in rheumatology* 27.6 (2015): 563-570.

van den Hoogen, Frank, et al. "2013 classification criteria for systemic sclerosis: an American College of Rheumatology/European League against Rheumatism collaborative initiative." *Arthritis & Rheumatism* 65.11 (2013): 2737-2747.

Volkmann, Elizabeth R., Kristofer Andréasson, and Vanessa Smith. "Systemic sclerosis." *The Lancet* 401.10373 (2023): 304-318.

Volkmann, Elizabeth R., Donald P. Tashkin, and Vanessa Smith. "Re-envisioning clinical trial design in systemic sclerosis." *Expert Review of Clinical Immunology* 21.5 (2025): 555-565.

Zulian, Francesco, et al. "The Pediatric Rheumatology European Society/American College of Rheumatology/European League against rheumatism provisional classification criteria for juvenile systemic sclerosis." *Arthritis Care & Research* 57.2 (2007): 203-212.

689 # List of abbreviations

| Abbreviation | Definition |
|---|---|
| ACA | Anticentromere antibodies |
| ACR | American College of Rheumatology |
| ACR-CRISS | American College of Rheumatology Composite Response Index in Systemic Sclerosis |
| ADA | Anti-drug antibody |
| AE | Adverse event |
| ANA | Antinuclear antibodies |
| CGA | Clinician Global Assessment |
| CI | Confidence interval |
| COPD | Chronic obstructive pulmonary disease |
| dcSSc | Diffuse cutaneous SSc |
| DLCO | Diffusing capacity of the lungs for carbon monoxide |
| EMA | European Medicines Agency |
| EULAR | European alliance of associations for rheumatology |
| FACIT-Fatigue | Functional Assessment of Chronic Illness Therapy–Fatigue |
| FEV | Forced Expiratory Volume |
| FVC | Forced vital capacity |
| GI | Gastrointestinal |
| HAQ-DI | Health assessment questionnaire–disability index |
| HRCT | High resolution computer tomography |
| ICE | Intercurrent events |
| ICH | International Council for Harmonisation |
| ILD | Interstitial lung disease |
| J4S | Juvenile Systemic Sclerosis Severity Score |
| jSSc | Juvenile systemic sclerosis |
| lcSSc | Limited cutaneous SSc |
| MIDD | Model Informed Drug Development |
| mRSS | modified Rodnan Skin Score |

| Abbreviation | Definition |
|---|---|
| nAb | Neutralising antibody |
| NRS | Numeric rating scale |
| PAH | Pulmonary arterial hypertension |
| PGA | Patient Global Assessment |
| PD | Pharmacodynamics |
| PK | Pharmacokinetics |
| PRO | Patient Reported Outcome |
| rCRISS | revised CRISS |
| RCS | Raynaud Condition Score |
| SGRP | St George's Respiratory Questionnaire |
| SHAQ | Scleroderma health assessment questionnaire |
| SSc | Systemic sclerosis |
| UCLA SCTC GIT | University of California Los Angeles Scleroderma Clinical Trial Consortium Gastrointestinal Tract |
| VAS | Visual analogue scale |
| VEGF | Vascular endothelial growth factor |
| WHO | World Health Organisation |

690