nature reviews rheumatology

https://doi.org/10.1038/s41584-023-00993-7

Consensus statement

Check for updates

Towards a novel clinical outcome assessment for systemic lupus erythematosus: first outcomes of an international taskforce

Kathryn Connelly[1] ✉, Laura E. Eades[1], Rachel Koelmeyer[1], Darshini Ayton[2], Vera Golder[1], Rangi Kandane-Rathnayake[1], Kate Gregory[1], Hermine Brunner[3], Laurie Burke[4], Laurent Arnaud[5], Anca Askanase[6], Cynthia Aranow[7], Ed Vital[8,9], Guillermo Pons-Estel[10], Khadija Dantata[11], Jeanette Andersen[12], Alain Cornet[12], Joy Buie[11], Ying Sun[13], Yoshiya Tanaka[14], Lee Simon[15], Youmna Lahoud[16], Alan Friedman[17], Kenneth Kalunian[18], Qing Zuraw[19], Victoria Werth[20], Sandra Garces[21], Eric F. Morand[1] & the TRM-SLE Consortium*

Abstract

Systemic lupus erythematosus (SLE) is a disease of high unmet therapeutic need. The challenge of accurately measuring clinically meaningful responses to treatment has hindered progress towards positive outcomes in SLE trials, impeding the approval of potential new therapies. Current primary end points used in SLE trials are based on legacy disease activity measures that were neither specifically designed for the clinical trial context, nor developed according to contemporary recommendations for clinical outcome assessments (COAs), such as that substantial patient input should be incorporated into their design. The Treatment Response Measure for SLE (TRM-SLE) Taskforce is a global collaboration of SLE clinician–academics, patients and patient representatives, industry partners and regulatory experts, established to realize the goal of developing a new COA for SLE clinical trials. The aim of this project is a novel COA designed specifically to measure treatment effects that are clinically meaningful to patients and clinicians, and intended for implementation in a trial end point that supports regulatory approval of novel therapeutic agents in SLE. This Consensus Statement reports the first outcomes of the TRM-SLE project, including a structured process for TRM-SLE development.

A full list of affiliations appears at the end of the paper. *A list of authors and their affiliations appears at the end of the paper. ✉e-mail: kathryn.connelly@monash.edu

Sections

Introduction

Systemic lupus erythematosus (SLE) is a chronic autoimmune disease characterized by multi-system involvement and unpredictable fluctuations in disease activity. For a substantial proportion of affected patients, current therapeutic strategies are insufficient[1]. Uncontrolled disease activity driven by autoimmunity combines with unwanted consequences of therapy to contribute to irreversible organ damage, the accumulation of comorbidities, and negative effects on patients' lives. Individuals with SLE can experience numerous severe symptoms, impaired quality of life and reduced function[2], and the disease is one of the leading causes of death in young women[3].

Despite recognition of the unmet therapeutic need and the identification of many promising drug targets in SLE, late-phase clinical trial successes and regulatory approvals of novel treatments have been few and far between. Although critical review of the contributory factors has led to some evolution in trial design[4], concerns about the inconsistent performance of SLE trial end points and how best to establish the efficacy of new therapies remain unresolved. In particular, the most common efficacy end points in current use, the SLE Responder Index (SRI) and the British Isles Lupus Assessment Group (BILAG)-based Composite Lupus Assessment (BICLA) both have well-recognized limitations[5]. Consequently, the development of new, evidence-based and discriminatory outcome measures to determine treatment effect in clinical trials is a major research priority in SLE.

The Treatment Response Measure for SLE (TRM-SLE) Taskforce was established to execute a project to develop and subsequently validate a novel clinical outcome assessment (COA) specifically intended for implementation as a primary outcome measure in SLE clinical trials that support regulatory approval of therapeutic agents. In this Consensus Statement we describe the first consensus outcome of the project that defines the high-level measurement goals that will underpin the development of the TRM-SLE COA, as well as a consensus on the research methods that will lead to an operational COA that can be incorporated into and validated in future clinical trials.

Table 1 | Governance committees comprising the TRM-SLE Taskforce

Committee	Members	Role
Steering Committee	35 members: Core research team (principal investigator, clinician–researchers and fellows, methodology experts) Patient and/or patient-organization representatives Industry representatives Metrology and regulatory experts Additional clinician–researchers with relevant expertise	Responsible for the scientific direction and delivery of the project
Scientific Advisory Board	29 members: Clinician–researchers with relevant expertise Patient and/or patient-organization representatives Industry representatives	Provide additional scientific input and oversight
Patient Advisory Panel	16 members: Patients with SLE Patient-organization representatives (Lupus Foundation of America and Lupus Europe)	Provide guidance regarding the patient experience of SLE via direct involvement and oversight of instrument development
Industry Advisory Board	28 members: Representatives from 10 industry partners with relevant expertise	Advise on industry-specific matters including leading regulatory engagement

SLE, systemic lupus erythematosus; TRM-SLE, treatment response measure for SLE.

Methods

Composition of the TRM-SLE Taskforce

The TRM-SLE Taskforce (also referred to herein as the TRM-SLE Consortium) was established in January 2022 explicitly to fulfil the goal of developing a new COA for use in SLE clinical trials. Acknowledging the need for input from multiple stakeholders, the taskforce consists of four governance committees (a Steering Committee, Scientific Advisory Board, Patient Advisory Panel and Industry Advisory Board). Potential taskforce members were nominated by the core research team and industry partners on the basis of the following criteria: clinical expertise in SLE, demonstrable experience in SLE clinical trials (for example as a principal investigator), expertise in outcome measurement and/or other relevant methodological expertise. Experts were directly approached and invited to participate. Patient and patient representative members were suggested by SLE patient organizations, including the Lupus Foundation of America and Lupus Europe, with the intention to recruit patients with clinical research and prior advocacy experience. We endeavoured to achieve representation of all major geographical regions for both clinician and patient representatives. Industry representatives were nominated by each of the collaborating pharmaceutical companies, and all companies known to the investigators to be active in SLE drug development were approached. The taskforce currently consists of 78 members in total. The TRM-SLE governance structure and committee roles are summarized in Table 1.

Generation of the study protocol for instrument development (Stage 1.1)

The overall TRM-SLE project is divided into stages of instrument development (Stage 1) followed by instrument validation (Stages 2 and 3), as illustrated in Fig. 1. Instrument development, which is the focus of this Consensus Statement, is divided into five sub-stages (Stages 1.1–1.5). In Stage 1.1, a detailed protocol was developed describing the specific methods to define the high-level measurement goals and context of use for the TRM-SLE instrument (Stage 1.2), to select the domains to be included in the TRM-SLE instrument (Stage 1.3), to determine how they will be measured (Stage 1.4), and to determine how domains will be incorporated into an overall definition of treatment response (Stage 1.5).

The study protocol was initially drafted by the core research team on the TRM-SLE Steering Committee (K. Connelly, L.E., R.K., D.A., V.G., R.K-.R. and E.M.). Protocol development was informed by an extensive review of the literature[5], focusing on understanding the characteristics and limitations of current SLE clinical trial end points, the complexities of outcome measurement in SLE and potential strategies to overcome these limitations via new approaches. Stages in TRM-SLE instrument development were also designed to align with the steps recommended by the relevant FDA Patient-Focussed Drug Development Guidance and related guidance from the Professional Society for Health Economics and Outcomes Research[6,7].

The draft protocol was circulated electronically to all members of the TRM-SLE Taskforce and initial feedback provided by e-mail correspondence. A Protocol Working Group was then established, consisting of a subset of 22 volunteering members of the TRM-SLE Taskforce, to

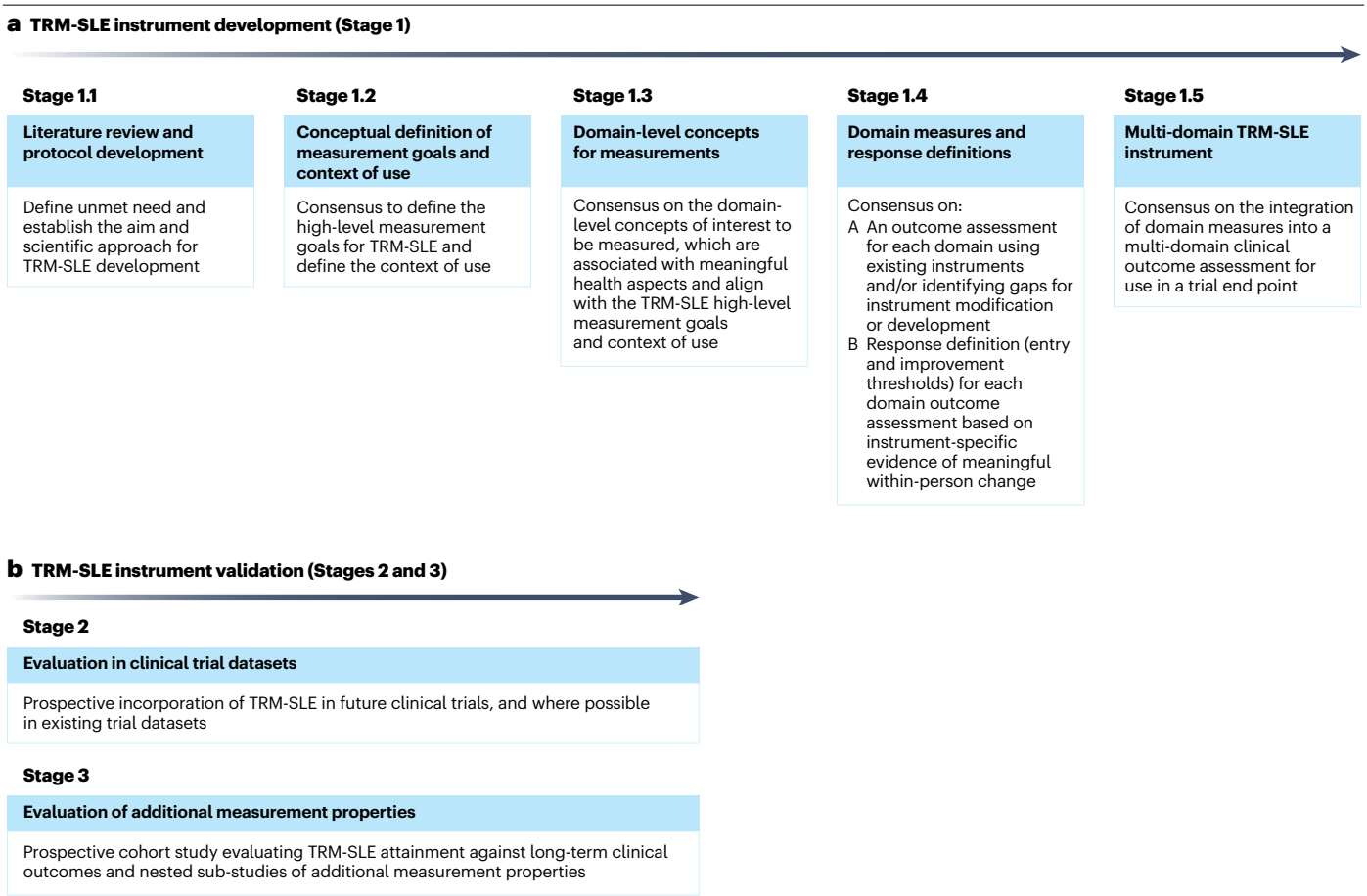

Fig. 1 | Steps in the development and validation of a novel treatment response measure for SLE. a, Stage 1 (instrument development) in the definition of a novel clinical outcome assessment via expert consensus informed by available data comprises five sub-stages that will lead to an operational, multi-domain, clinical outcome assessment that defines treatment response at both a global and a domain-specific level. The outcomes of Stages 1.1–1.2, and consensus methods agreed upon by the taskforce for Stages 1.3–1.5, are described in this Consensus Statement. **b**, The provisional instrument will be validated in trial datasets (Stage 2) as well as concurrently undergoing additional testing of its measurement properties (Stage 3). Key measurement properties that will be evaluated include construct validity, reliability, ability to detect change, discrimination of treatment effects, interpretability and feasibility.

work through proposed changes to the draft protocol. The Protocol Working Group (K. Connelly, L.E., R.K., D.A., H.B., L.B., L.A., A.A., C.A., E.V., G.P.-E., K.D., Y.S., Y.T., L.S., Y.L., A.F., K.K., Q.Z., V.W., S. Garces and E.M.) included experts in COA development, consensus methodology and regulatory affairs, industry partners, patient representatives and SLE clinician–researchers. The Protocol Working Group revised the methodological steps via group discussion over three virtual meetings and e-mail correspondence. During protocol revision, the proposed methods were also presented at a meeting of the Patient Advisory Panel, and patient feedback was collected and incorporated into changes made by the Protocol Working Group. A revised protocol document was prepared by the core research team, circulated electronically and individually approved by TRM-SLE Taskforce members.

Methods to define high-level measurement goals for TRM-SLE development (Stage 1.2)

Following protocol development, the first major scientific step towards instrument development was to achieve consensus on the high-level goals of measurement underpinning the TRM-SLE project (Stage 1.2). This process occurred in two parts: first, the measurement goals of TRM-SLE were conceptually defined using the Patients/Population, Intervention, Comparator/Control, Outcome/Objective, Context (PICO-C) framework; and second, consensus was established on the context of use for the TRM-SLE instrument. From a regulatory perspective, the context of use is a specific statement describing the manner and purpose of use of an instrument, including a description of the targeted disease and study population, study design and setting for its intended use[6,7]. Defining the context of use is a key step outlined in regulatory guidance pertaining to COA development, to help to ensure that any new instrument is developed and validated for its intended scope of implementation, so that results can be appropriately interpreted and applied[6,7].

Members of the TRM-SLE Steering Committee and Scientific Advisory Boards, encompassing representation of key stakeholder groups, were invited to participate in the definition of the high-level measurement goals and context of use for TRM-SLE. Participants (A.A., A.C., A.F., C.A., C.B., C. Sibley., C. Stach., D.A., E.V., E.M., E.Z., G.S., G.P.-E., H.A., H.B., J.A., J.B., J. Merrill, J. Maller, J.R.T., K. Costenbader, K.

Box 1

Key limitations of outcome measures currently used to determine treatment response in SLE clinical trials

- Lack of patient input in determining which concepts are important to be measured.
- Use of discrete thresholds introducing floor and ceiling effects and limiting the ability to capture variations in the severity of manifestations.
- Numerical thresholds defining improvement and fixed weightings applied to manifestations that are not based on empirical evidence of meaningful within-patient change.
- Developed using post hoc analysis of trial data, risking bias to specific drug mechanisms or specific study designs.
- Some included manifestations are poorly defined and/or rarely appear in clinical trials.
- Adoption in clinical trials prior to extensive testing of measurement properties.
- Can be complex and non-intuitive to complete and interpret.

Connelly, K.D., K.G., K.K., L.E., L.A., L.B., M.D., M.S., M.M., N.D., P.M., Q.Z., R.K., R.K.-R., R.F., R.v.V., S. B., S. Garces, V.G., V.W., Y.S., Y.T. and Y.L.) met over a series of four virtual meetings (three meetings addressing the high-level measurement goals and one meeting addressing the context of use). The participation in each meeting is indicated in Supplementary Table 1. Item generation was facilitated using a web-based application (MURAL; https://www.mural.co/), which enables participants to post ideas via 'sticky-notes' onto a common virtual whiteboard in real time. Suggested items were then grouped and refined via moderated discussion in an interactive fashion. Online polling integrated into the virtual meeting platform was used for voting on the proposed final wording of each element of the conceptual definition and context of use, with a pre-defined consensus threshold set at 70% agreement. Only members present in the meetings participated in voting, and voting was not compulsory (to enable committee members without SLE-specific expertise to opt out of voting at their discretion). Consensus statements and percentage agreement were recorded for each element of the final conceptual definition and context of use. The consensus outcomes were then circulated electronically and presented at virtual meetings of each TRM-SLE governance committee (Table 1), for final approval.

Rationale for development of TRM-SLE

Although some recent positive outcomes have been achieved in SLE clinical trials, including the approval of anifrolumab (an antibody that targets the type I interferon receptor), inconsistent end point performance continues to affect the accurate interpretation of the treatment effects of novel agents. This inconsistency is demonstrated by positive phase II trial results of drug candidates that fail to be replicated in subsequent phase III trials[8–12], phase III trials with identical study protocols that produce conflicting results with regard to their primary end points[11–14], and discrepancies between the outcomes of primary study end points and other clinically relevant measures, such as steroid-sparing effects[15]. Recent examples of these challenges include the phase III clinical trials of anifrolumab (TULIP studies), the Janus kinase inhibitor baricitinib (BRAVE studies) and ustekinumab, a monoclonal antibody that blocks the p40 subunit shared by IL-12 and IL-23, each of which followed on from successful phase II studies. In the TULIP-1 and TULIP-2 phase III trials of anifrolumab that used the same eligibility criteria, TULIP-1 found no significant difference between anifrolumab and placebo using the SRI-4 as the primary end point, but differences favouring anifrolumab were detected using BICLA as an alternative measure of overall efficacy[13]. By contrast, TULIP-2 detected a significant treatment effect of anifrolumab using BICLA as a primary end point, as well as SRI-4 as a secondary end point[14]. Both studies also met other key secondary efficacy end points, ultimately resulting in regulatory approval of anifrolumab in several countries. Similar to the TULIP studies, the recent SLE-BRAVE-I and SLE-BRAVE-II phase III studies of baricitinib also produced conflicting results, meeting the SRI-4 primary end point in BRAVE-I, but failing to meet the same primary end point in BRAVE-II[11,12]. Meanwhile, phase III trials of ustekinumab were abandoned because of negative results in an interim analysis, again following promising results in phase II testing[10]. Inconsistency between results with different end points within the same population are also characteristic of some of these trials[16], and of a trial of belimumab in childhood SLE[17].

The COAs currently used for determination of the treatment response in SLE trials are imperfect; we have previously reviewed the factors behind their inconsistent performance[5], and summarize their key limitations in Box 1. Many issues stem from the fact that COAs incorporated in current trial end points, such as the SLE Disease Activity Index (SLEDAI) and BILAG were primarily developed as disease activity measures and were repurposed for the measurement of treatment response in the absence of better alternatives. As a consequence, the concepts measured and thresholds used for defining meaningful improvement were not grounded in the context of clinical trial use, and importantly, lacked substantial patient input. Current recommendations for COAs intended for use in clinical trials highlight the importance of these instruments being sufficiently validated for their specific scope of intended use, and the vital role that the patient perspective has in ensuring that the interpretation of outcomes reflects meaningful health aspects[12,13]. Therefore, COAs used in current SLE trial primary end points do not meet contemporary measurement standards, which in combination with their history of unreliable performance, underpins the major need for new instruments for this purpose. Such new instruments should specifically seek to avoid replicating the liabilities of legacy measures and follow modern recommendations for instrument development, including guidance documents published by regulatory bodies such as the FDA and EMA[6,7].

Protocol for TRM-SLE instrument development

The first consensus outcomes relating to TRM-SLE instrument development are described in detail below. The first two stages (Stages 1.1 and 1.2) have been completed, yielding consensus on the high-level measurement goals and context of use for TRM-SLE and a detailed protocol describing specific methods for future stages of instrument development (Stages 1.3–1.5). Terminology related to COA development used

in FDA and Professional Society for Health Economics and Outcomes Research guidance documents[6,7] was adopted in the study protocol and throughout this Consensus Statement.

Conceptual framework for TRM-SLE

The recommended steps towards a fit-for-purpose COA can be visually depicted in the form of a conceptual framework (Fig. 2) that summarizes the following key elements: relevant health outcomes (symptoms, signs and effects of the disease) in the target population; specific concepts of interest targeted for assessment; COAs proposed to measure each concept of interest, potentially including existing, modified and novel COAs; and the use of these COAs to generate a score for each concept and to define treatment response[10].

Patient involvement in TRM-SLE development

A key goal of TRM-SLE development is to ensure that the resultant COA captures aspects of health that are meaningful to both the patient and clinician. For this reason, the study protocol has been designed to incorporate the perspectives of both of these stakeholder groups, as illustrated in Fig. 3.

Consensus high-level measurement goals and context of use for TRM-SLE (Stage 1.2)

A panel of 45 taskforce members representative of the key stakeholder groups generated a conceptual definition for the high-level measurement goals of TRM-SLE using the PICO-C (Patients/Population, Intervention, Comparator/Control, Outcome, Context) framework. Over the course of three virtual meetings, multiple rounds of item generation, moderated discussion and real-time voting resulted in the conceptual definition for TRM-SLE, which overall achieved predefined levels of agreement (Table 2). Dissenting opinions were mostly with regard to a desire to expand beyond the traditional study design of SLE clinical trials, for example, by consideration of the patient populations to whom TRM-SLE should apply. Although most participants expressed a preference to focus on active immune-mediated disease manifestations for the purpose of a clinical trial, others suggested additional

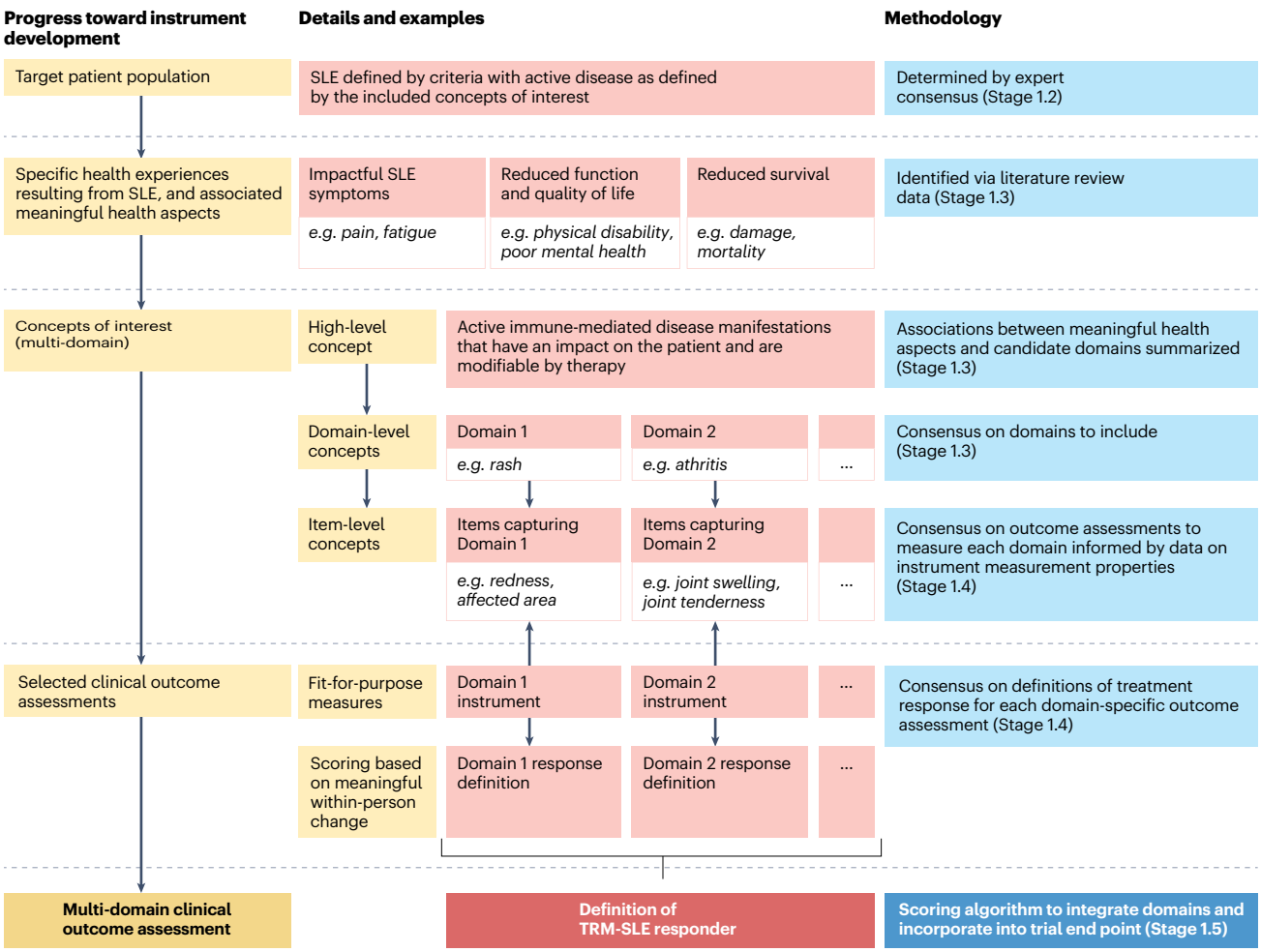

Fig. 2 | Conceptual framework for TRM-SLE. Illustration of the conceptual framework that will be established using the proposed methods for development of the treatment response measure for systemic lupus erythematosus (TRM-SLE) instrument. As recommended by regulatory guidance, this framework includes the concepts of interest (including domain-level and item-level concepts) targeted for assessment and how these relate to the patient experience, along with consideration of how the concepts will be measured and scored. In this figure, italic text represent theoretical examples; the specific elements that will fill the framework will be determined by the completion of Stages 1.3–1.5.

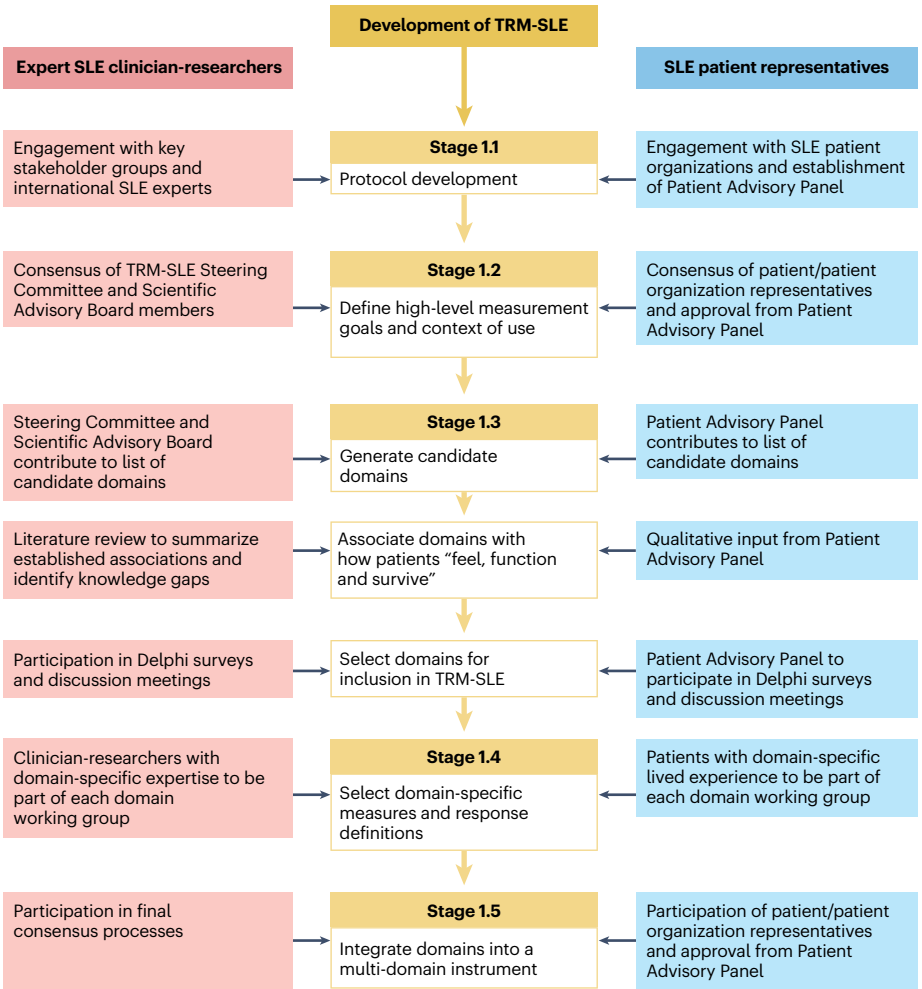

Fig. 3 | Input of SLE patient representatives and expert clinician–researchers in TRM-SLE development. An illustration of the planned involvement of patient partners and expert systemic lupus erythematosus (SLE) clinician–researchers in the development process for the treatment response measure for systemic lupus erythematosus (TRM-SLE).

populations with unmet therapeutic needs in SLE, or the ability to assess interventions targeting non-inflammatory disease features, or use in settings beyond randomized controlled trials. These suggestions were balanced against the regulatory requirement for COA approval to have a pre-specified and very specific context of use. With reference to the consensus PICO-C framework, a statement was derived describing the high-level concept that should be captured by the TRM-SLE instrument as "active immune-mediated disease manifestations that impact the patient and are modifiable by therapy to reduce or control disease activity" (Box 2).

Expanding upon this high-level conceptual definition, consensus on the primary context of use for which TRM-SLE will be developed was sought. Over the course of a single virtual meeting, moderated discussion and voting focussed on defining the target disease and study subpopulation, study setting and trial design for which the TRM-SLE instrument will be primarily developed. This resulted in the context of use detailed in Table 3, which achieved 100% consensus among 30 participating taskforce members. We anticipate that the described primary context of use may be refined in an iterative manner based on the outcomes of the subsequent steps in instrument development, which includes the incorporation of TRM-SLE into an end point, and its associated analysis plan.

Future stages of TRM-SLE development

The methods described below represent an approach that has both similarities to and differences from processes adopted in rheumatic disease outcome measurement more generally. Given the goal of developing an outcome measure suitable for supporting regulatory approval of new agents in clinical trials, particular attention has been paid to ensuring that methods conform to published recommendations that apply to this specific context[6,7], while accounting for disease-specific measurement challenges and the intention to avoid replicating known limitations of current SLE trial outcome measures.

Selection of domains to be measured in TRM-SLE (Stage 1.3)

The next stage of TRM-SLE development will select domain-level concepts to be measured by the instrument. A concept (also known as a concept of interest), for regulatory purposes, is a health aspect that is intended to be captured by a COA. Current guidance emphasizes that these aspects of health should be meaningful for patients, defined as having an effect on how the patient "feels, functions or survives"[6]. As illustrated in the pyramid within Fig. 2, concepts can be considered at multiple levels. The high-level concept to be measured by TRM-SLE will be captured by consideration of multiple sub-concepts or 'domains'. These domains might include organ-based and/or system-based

Glossary

Clinical outcome assessment
Assessment of a clinical outcome that describes or reflects an aspect of health, and that can be made through a report by a clinician (clinician-reported outcome measures), a patient (patient-reported outcome), a non-clinician observer (observer-reported outcome), or through a performance-based assessment (performance outcome).

Concepts of interest
In a regulatory context, the aspect of an individual's clinical, biological, physical or functional state or experience that the assessment is intended to capture (or reflect).

Context of use
A statement that fully and clearly describes the way the outcome assessment is to be used and the regulated product-development purpose.

Domain
A sub-concept represented by a score of an instrument that measures a larger concept comprising multiple domains.

End points
Precisely defined variables intended to reflect an outcome of interest that is statistically analysed to address a particular research question, including the type and timing of assessments, the assessment tools used, and other details, as applicable, such as how multiple assessments within an individual are to be combined.

Fit-for-purpose
A conclusion that the level of validation associated with a clinical outcome assessment is sufficient to support its context of use.

Meaningful health aspects
Aspects of health (feelings, functions or survival) adversely affected by the disease, which the patient cares about and has a preference that they do not become worse, or that they improve, or that they are prevented.

manifestations (such as lupus nephritis), symptoms of SLE (such as rash), or in some cases manifestations measured by laboratory tests or other investigations (such as thrombocytopenia). Each domain will then be defined by one or more 'items', which are the specific concepts measured and scored to produce a representation of each domain. For example, a hypothetical domain that falls under the high-level concept for TRM-SLE could be 'arthritis'. The items assessed to capture this domain potentially include concepts such as joint swelling, tenderness or pain. The focus of Stage 1.3 is the selection of domains to be captured in the TRM-SLE instrument. The specific COA(s) with which each domain is measured, and the associated item-level concepts, are addressed in Stage 1.4.

Generation of candidate domains and associations with meaningful health aspects. Clinician–researchers with lupus expertise, industry representatives with SLE trial experience, and patients will nominate candidate domains. These nominations will be grouped and refined to produce a core list of domains to be rated for inclusion in TRM-SLE.

The TRM-SLE instrument is intended to be a patient-centred measure, so the domains considered in the determination of response to therapy must be associated with health effects that are meaningful from the patient perspective (meaningful health aspects), including symptoms and functional effects identified to be important by patients themselves, and associations with outcomes of prognostic importance such as damage and mortality. Evidence of associations of the core list of candidate domains with meaningful health aspects will be evaluated by a targeted review of the literature, and summarized to inform domain selection during a modified Delphi process, as described below. Complementary to the literature review, the Patient Advisory Panel will also provide patient perspectives on the candidate domains and their associations with meaningful health aspects, and help to prioritize future research agendas where evidence gaps are identified.

Selection of domains for inclusion in TRM-SLE: modified Delphi process. Consensus on which of the candidate domains will be included in the TRM-SLE instrument will be achieved by a two-part modified Delphi process. In the first part, domains will be rated on their 'importance', defined as the extent to which a domain is associated with meaningful health aspects (impact on how a person with SLE "feels, functions or survives"). Clinicians with expertise in SLE clinical care and trials, including experts external to the TRM-SLE project, and patient representatives via the Patient Advisory Panel, will participate in two to three Delphi survey rounds, with discussion meetings between voting rounds. We plan to recruit a total of 50–100 participants, with international representation. Participants will rate each domain for 'importance' on a nine-point scale, with ratings assigned to three categories: ≥7 (critically important to include), 4–6 (important but not critical) and ≤3 (not important)[18]. A summary of evidence from the literature associating candidate domains with meaningful health aspects will be provided to inform participant ratings and support panel discussions.

Consensus on domain 'importance' will be defined as ≥70% of total participants scoring 7–9 (ref. 14) with the additional requirement that the consensus threshold is met in both expert clinician and patient groups. Domains achieving consensus on 'importance' will proceed to a second set of ratings, where participants will rate domains on three additional characteristics relevant to their inclusion. The first characteristic is 'appropriateness', defined as whether the domain is an active immune-mediated-disease manifestation that is modifiable by therapy to reduce or control disease activity in an SLE clinical trial (as defined by the measurement goals and context of use for TRM-SLE arising from Stage 1.2). The second characteristic is 'representation', defined as whether domain activity occurs with sufficient frequency in patients

Table 2 | Conceptual definition of the high-level goals of measurement for TRM-SLE

PICO-C	Consensus definition	Number of contributors to definition	Agreement[a]
Patients or population	SLE defined by criteria, with active immune-mediated disease manifestations modifiable by therapy	41	81%
Intervention	Treatment to reduce or control disease activity	33	92%
Comparator or control	Placebo and/or active comparator	33	96%
Outcome or objective	The impact of an intervention on the patient, as measured by change in the concepts of interest	33	100%
Context	Clinical trials assessing efficacy and satisfying requirements for registration	33	96%

[a]Percentage agreement for each consensus definition amongst voting taskforce members.
PICO-C, Patients/Population, Intervention, Comparator/Control, Outcome, Context.

Box 2

High-level concept to be measured by the TRM-SLE

"Active immune-mediated disease manifestations that impact on the patient and are modifiable by therapy to reduce or control disease activity".

with SLE and active disease to warrant inclusion in TRM-SLE. The third characteristic is 'measurability', which is defined as whether the domain can be clearly defined and treatment response accurately quantified. Domains must achieve ≥70% of participants scoring ≥7 on all three of these additional characteristics to meet consensus for inclusion in the TRM-SLE instrument.

Domain measures and response definitions (Stage 1.4)
Once consensus has been achieved on TRM-SLE domains, individual working groups will be established for each domain to achieve consensus on selection of an outcome assessment to measure the domain and/or outline a necessary research agenda if a fit-for-purpose measure is not available, and on how to numerically define domain-specific responses using the chosen outcome assessment, based on entry and improvement thresholds that are anchored to evidence of meaningful within-person change.

Domain-specific working groups. Each domain working group will comprise 6–12 individuals with expertise relevant to their specific allocated domain. The composition of each working group may vary depending on the nature of the domain, but will include clinician–researchers with domain-specific clinical, trial and/or measurement expertise and patient representatives with lived experience of the affected SLE domain.

Selection of domain measures. Members of each working group will nominate ideas for suitable candidate instruments to measure their target domain. A nominal group technique will then be used to rank and achieve consensus on a preferred domain-specific outcome assessment, which will be followed by a systematic literature review of the measurement properties of the selected outcome assessment for each domain, particularly as they pertain to the TRM-SLE context of use, to determine whether the instrument is fit for purpose or whether additional studies are required. At the discretion of the working group, more than one candidate instrument could proceed to the review step, with a final decision incorporating evidence of the measurement properties of the candidate instruments. Evaluation of measurement properties will include consideration of face validity (working group expert opinion), content validity (evidence that the included items adequately capture the domain, and that the instrument score represents the intended measurement concept), feasibility (working group expert opinion and evidence from use in previous SLE clinical trials), reliability (evidence of test–retest, intra-rater and inter-rater reliability), construct validity (evidence of associations with other relevant measures, through analysis of cohort and clinical trial datasets), ability to detect change (evidence of responsiveness to change in the concept being measured), meaningful thresholds and interpretability (evidence of thresholds of meaningful disease activity and improvement anchored to appropriate patient-centred outcomes), and discrimination of a treatment effect (evidence from previous SLE clinical trials). If further instrument validation is required, it will be planned in subsequent validation stages using existing trial datasets or as an exploratory end point in the setting of a new trial.

Definition of domain-specific response
Once an outcome assessment has been identified for the measurement of a particular domain, the working group will establish consensus on the numerical definition of response for that domain. For each domain-specific outcome assessment, a response definition will have two components. The first will be the entry threshold: the minimum level of domain-specific disease activity at study entry, from which improvement can be measured (reflecting a level of severity associated with a meaningful clinical effect and permitting sufficient room for improvement). The second component will be the improvement threshold: the minimum level of improvement (from baseline) required to be defined as a responder in a specific domain (reflecting a clinically important within-person change anchored to appropriate patient-centred outcomes, and sufficiently stringent to discriminate a treatment effect between arms in a clinical trial). Similar to the process to select domain measures, members of the domain working group will nominate candidate response definitions. This process will be informed by available empirical data using anchor-based methods supporting specific thresholds of activity and improvement associated with meaningful clinical effect and clinical benefit, respectively. A nominal group technique will then be used to rank and achieve consensus on a particular response definition to be included in the final TRM-SLE instrument, subject to further validation in subsequent trial datasets where required.

Integration of domain measures into a multi-domain COA (Stage 1.5)
Once consensus is achieved on which concepts will be measured by the TRM-SLE instrument, the outcome assessments that will be used for measurement of these concepts, and the thresholds that define meaningful within-patient change for each domain, the final stage of instrument development will establish consensus on how to integrate these measures into a multi-domain COA that can be interpreted to define treatment responders when deployed as part of a trial end point in validation studies. Specific points for consensus at this stage will include defining the scoring algorithm that specifies an overall responder (including possible weighting of different domains), and determining methods to capture worsening or new activity that develops over the course of a trial (such as in domains that are not specifically measured in the TRM-SLE instrument).

Validation of the TRM-SLE instrument (Stages 2 and 3)
Following instrument development, the taskforce plans an extensive validation programme to evaluate the performance of the multi-domain TRM-SLE measure and to ensure that scoring and definition of treatment response within a trial end point reflect meaningful within-patient change. Validation will include prospective incorporation of TRM-SLE as an exploratory end point in future clinical trials, as well as evaluation of TRM-SLE in available existing trial datasets, where possible. Key measurement properties, including construct validity, the ability to detect change and the ability to discriminate a treatment effect (for example, between treatment arms in a clinical trial) will be assessed,

Table 3 | Primary context of use for the TRM-SLE instrument

Context of use	Description	Agreement[b]
Targeted disease and study subpopulation	Adult and adolescent patients with criteria-defined SLE, with active disease in one or more included domains[a] despite standard-of-care therapy	100%
Targeted study design and study setting	Randomized controlled trial conducted in an outpatient setting, with the TRM-SLE instrument being the main outcome assessment within a primary efficacy end point with landmark analysis comparing meaningful response between treatment arms	100%

[a]Expanded definition: active disease in one or more included domains with sufficient domain-specific activity to have a meaningful impact on the patient (domains confirmed in Stage 1.3 and sufficient domain-specific activity numerically defined in Stage 1.4 using chosen domain measurement instrument). [b]Percentage agreement for each component of the context of use amongst voting members of the 30 taskforce members present.

including comparison with legacy disease activity measures used for trial eligibility (such as SLEDAI and BILAG) and with currently used responder indices (such as SRI and BICLA). This process will also enable evaluation of operational considerations affecting implementation of TRM-SLE as a trial end point in its defined context of use. Concurrently, it is intended to conduct a prospective cohort study to validate attainment of TRM-SLE response against long-term clinical outcomes, including patient-reported outcomes, damage accrual, flare and attainment of target disease activity states such as Lupus Low Disease Activity State[19] and remission[20]. Studies addressing other measurement properties, including reliability, are planned via sub-studies nested within the prospective cohort study, along with additional case-based studies.

Conclusions

This Consensus Statement reports the first consensus outcomes and agreed study protocol of an international taskforce specifically established to develop a novel COA for SLE trials. This new outcome measure will be specifically designed for the clinical trial context, will learn from the limitations of legacy trial outcome measures, will follow updated regulatory and instrument development guidance, and will incorporate input from all key stakeholder groups, importantly including the patient voice. The new outcome measure has the potential to be used as a composite measure, but also as a source of individual domain measures, and in contexts additional to the defined context of use. Although a challenging endeavour, it is the ambition of the TRM-SLE project to develop an outcome measure in SLE that is accepted by key stakeholders (including regulators), that is successfully incorporated into future registration trials, and that leads to clearer interpretation of the efficacy of new treatments that reflect both patient and clinician health priorities.

Published online: 11 July 2023

References

1. Kandane-Rathnayake, R. et al. 'Not at target': prevalence and consequences of inadequate disease control in systemic lupus erythematosus — a multinational observational cohort study. *Arthritis Res. Ther.* **24**, 70 (2022).
2. Schmeding, A. & Schneider, M. Fatigue, health-related quality of life and other patient-reported outcomes in systemic lupus erythematosus. *Best. Pract. Res. Clin. Rheumatol.* **27**, 363–375 (2013).
3. Yen, E. Y. & Singh, R. R. Brief report: lupus — an unrecognized leading cause of death in young females: a population-based study using nationwide death certificates, 2000–2015. *Arthritis Rheumatol.* **70**, 1251–1255 (2018).
4. Merrill, J. T. et al. Lupus community panel proposals for optimising clinical trials: 2018. *Lupus Sci. Med.* **5**, e000258 (2018).
5. Connelly, K., Golder, V., Kandane-Rathnayake, R. & Morand, E. Clinician-reported outcome measures in lupus trials: a problem worth solving. *Lancet Rheumatol.* **3**, e595–e603 (2021).
6. US Department of Health and Human Services Food and Drug Administration. Patient-Focused Drug Development Guidance Series for Enhancing the Incorporation of the Patient's Voice in Medical Product Development and Regulatory Decision Making. Selecting, Developing, or Modifying Fit-for-Purpose Clinical Outcome Assessments (Draft Guidance). https://www.fda.gov/media/159500/download (2022).
7. Powers, J. H. 3rd et al. Clinician-reported outcome assessments of treatment benefit: report of the ISPOR clinical outcome assessment emerging good practices task force. *Value Health* **20**, 2–14 (2017).
8. van Vollenhoven, R. F. et al. Efficacy and safety of ustekinumab, an IL-12 and IL-23 inhibitor, in patients with active systemic lupus erythematosus: results of a multicentre, double-blind, phase 2, randomised, controlled study. *Lancet* **392**, 1330–1339 (2018).
9. Wallace, D. J. et al. Baricitinib for systemic lupus erythematosus: a double-blind, randomised, placebo-controlled, phase 2 trial. *Lancet* **392**, 222–231 (2018).
10. van Vollenhoven, R. F. et al. Phase 3, multicentre, randomised, placebo-controlled study evaluating the efficacy and safety of ustekinumab in patients with systemic lupus erythematosus. *Ann. Rheum. Dis.* **81**, 1556–1563 (2022).
11. Morand, E. F. et al. Baricitinib for systemic lupus erythematosus: a double-blind, randomised, placebo-controlled, phase 3 trial (SLE-BRAVE-I). *Lancet* **401**, 1001–1010 (2023).
12. Petri, M. et al. Baricitinib for systemic lupus erythematosus: a double-blind, randomised, placebo-controlled, phase 3 trial (SLE-BRAVE-II). *Lancet* **401**, 1011–1019 (2023).
13. Furie, R. et al. Type I interferon inhibitor anifrolumab in active systemic lupus erythematosus (TULIP-1): a randomised, controlled, phase 3 trial. *Lancet Rheumatol.* **1**, E208–E219 (2019).
14. Morand, E. F. et al. Trial of anifrolumab in active systemic lupus erythematosus. *N. Engl. J. Med.* **382**, 211–221 (2020).
15. Oon, S., Huq, M., Godfrey, T. & Nikpour, M. Systematic review, and meta-analysis of steroid-sparing effect, of biologic agents in randomized, placebo-controlled phase 3 trials for systemic lupus erythematosus. *Semin. Arthritis Rheum.* **48**, 221–239 (2018).
16. Bruce, I. N. et al. Concordance and discordance in SLE clinical trial outcome measures: analysis of three anifrolumab phase 2/3 trials. *Ann. Rheum. Dis.* **81**, 962–969 (2022).
17. Brunner, H. I. et al. Safety and efficacy of intravenous belimumab in children with systemic lupus erythematosus: results from a randomised, placebo-controlled trial. *Ann. Rheum. Dis.* **79**, 1340–1348 (2020).
18. Maxwell, L. J. et al. Core domain set selection according to OMERACT filter 2.1: the OMERACT methodology. *J. Rheumatol.* **46**, 1014–1020 (2019).
19. Franklyn, K. et al. Definition and initial validation of a Lupus low disease activity state (LLDAS). *Ann. Rheum. Dis.* **75**, 1615–1621 (2016).
20. van Vollenhoven, R. F. et al. 2021 DORIS definition of remission in SLE: final recommendations from an international task force. *Lupus Sci. Med.* https://doi.org/10.1136/lupus-2021-000538 (2021).

Acknowledgements

K. Connelly is a recipient of a National Health and Medical Research Council (NHMRC) Postgraduate Research Scholarship and the Royal Australasian College of Physicians CRB Blackburn Top-Up Scholarship. E.M. is supported by an NHMRC Investigator Grant. The following ten companies have partnered with Monash University to participate in and financially support this project: AbbVie, Amgen, AstraZeneca, Biogen, Bristol-Myers Squibb, Eli Lilly, Genentech, Janssen, Merck KGaA and UCB.

Author contributions

K. Connelly, L.E. and E.M. researched data for the article. K. Connelly, L.E., R.K., D.A., V.G., R.K.-R., K.G., H.B. L.B., L.A., A.A., C.A., E.V., G.P.-E., K.D., J.A., A.C., J.B., Y.S., Y.T., L.S., Y.L., A.F., K.K., Q.Z., V.W., S. Garces and E.M. made a substantial contribution to discussion of the content. K. Connelly and E.M. wrote the article. All authors reviewed and/or edited the manuscript before submission.

Competing interests

H.B. has received grants from Pfizer and fees from Novartis, Pfizer and Roche. The Cincinnati Children's Hospital where H.B., the PRCSG Chairperson, works as a full-time employee, has received contributions for consultation from the following industries in the past 2 years: AbbVie, AstraZeneca-Medimmune, Biogen, Boehringer, Bristol-Myers Squibb, Celgene, Cerocor, Eli Lilly, F. Hoffmann-La Roche, Horizon, Idorsia, Janssen, Merck, Novartis and Sanofi. This funding has been reinvested for the research activities of the hospital in a fully independent manner, without any commitment to third parties. L.A. has received grants from GlaxoSmithKline and fees from AbbVie, Alexion, Amgen, AstraZeneca, Biogen, Boehringer-Ingelheim, Bristol-Myers Squibb, Eli Lilly, GlaxoSmithKline, Grifols, Janssen, LFB, Medac, Novartis, Pfizer, Chugai, Sêmeia and UCB. A.A. has received fees from AbbVie/Abbott, AstraZeneca, Amgen, Aurinia, Bristol-Myers Squibb, Eli Lilly, GlaxoSmithKline, Idorsia, Janssen, Pfizer and UCB. E.V. has received fees from Aurinia, Eli Lilly, Novartis, Pfizer, Roche and UCB. G.P.-E. has received fees from GlaxoSmithKline, Janssen, Novartis, Pfizer, Sanofi and Werfen/Inova. Y.T. has received grants from AbbVie, Asahi-Kasei, Boehringer-Ingelheim, Chugai, Daiichi-Sankyo, Eisai and Takeda, and fees from AbbVie, AstraZeneca, Boehringer-Ingelheim, Bristol-Myers Squibb, Chugai, Daiichi-Sankyo, Eisai, Eli Lilly, Gilead, GlaxoSmithKline, Mitsubishi-Tanabe and Pfizer. L.S. has received fees from Ampio, Amzell, Applied Biosystems, Aptinyx, AstraZeneca, Aura, Aurinia, Bexson, Biosplice, Boehringer-Ingelheim, Bone Med, Bone Therapeutics, Bristol-Myers Squibb, Cancer Prevention, Cerebral Therapeutics, Chemocentryx, Chemomab, Diffusion Bio, Direct, Eli Lilly, Elorac, EMD Serono, Enalare, Eupraxia, Foundry Therapeutics, Galapagos, Galvani, Gilead, GlaxoSmithKline, Histogen, Horizon Therapeutics, Idera, Ingel, Intravital, Kaniska, Kiel Labs, Kyoto, Mesoblast,

Minerva, Mpathix, Neema, Neumentum, Pfizer, PPD, Priovant, Protalix, Regenosine, Rigel, Roivant, Samus, Sana, SpineThera, StageBio, Theraly, TLC, Transcode, Unity, Urica, Viridian and Xalud. Y.L. is an employee of Biogen and owns shares in Biogen stock. A.F. is an employee of AbbVie and owns shares in AbbVie stock. K.K. has received fees from AbbVie/Abbott, Amgen, AstraZeneca, Aurinia, Biogen, Bristol-Myers Squibb, Eli Lilly, Equillium, Genentech, Gilead, Janssen, Nektar, Pfizer, Roche, Sanford Consortium and Viela. Q.Z. is an employee of Janssen. V.W. has received fees from AbbVie/Abbott, Amgen, AnaptysBio, Argenx, AstraZeneca, Biogen, Bristol Myers Squibb, Celgene, Corbus, CSL Behring, Eli Lilly, EMD Serono, Galderma, Genentech, Gilead, GlaxoSmithKline, Horizon Therapeutics, Janssen, Merck/MSD, Nektar, Novartis, Octapharma, Pfizer, Principia, Regeneron, Resolve, Rome Pharmaceuticals, Sanofi, Syntimmune and Viela. S. Garces is an employee of Amgen. E.M. has received grants from AbbVie, Amgen, AstraZeneca, Biogen, Bristol-Myers Squibb, Eli Lilly, EMD Serono, Genentech, GlaxoSmithKline, Janssen and UCB, and fees from Amgen, AstraZeneca, Biogen, Bristol-Myers Squibb, Capella, Eli Lilly, EMD Serono, Genentech, GlaxoSmithKline, Janssen, Neovacs, Novartis, Sanofi, Servier, UCB, Wolf and Zenas. All other authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41584-023-00993-7.

Peer review information *Nature Reviews Rheumatology* thanks F. Conti, W. Stohl and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

[1]School of Clinical Sciences at Monash Health, Monash University, Clayton, Victoria, Australia. [2]School of Public Health and Preventive Medicine, Monash University, Clayton, Victoria, Australia. [3]Division of Rheumatology, Cincinnati Children's Hospital, University of Cincinnati, Cincinatti, OH, USA. [4]LORA Group, Royal Oak, MI, USA. [5]Department of Rheumatology, National Reference Center for Autoimmune Diseases (RESO), Hopitaux Universitaires de Strasbourg, Strasbourg, Alsace, France. [6]Lupus Center, Columbia University Medical Center, New York City, NY, USA. [7]Lupus Centre of Excellence, Feinstein Institutes for Medical Research, Northwell Health, Manhasset, NY, USA. [8]Leeds Institute of Rheumatic and Musculoskeletal Medicine, University of Leeds, Leeds, UK. [9]NIHR Leeds Biomedical Research Centre, Leeds Teaching Hospitals NHS Trust, Leeds, UK. [10]Departamento de Medicina Interna, Centro Regional de Enfermedades Autoinmunes y Reumáticas de Grupo Oroño (GO CREAR), Rosario, Argentina. [11]Lupus Foundation of America, Washington, DC, USA. [12]Lupus Europe, Brussels, Belgium. [13]Merck Healthcare KGaA, Darmstadt, Hesse, Germany. [14]The First Department of Internal Medicine, University of Occupational and Environmental Health, Japan, Kitakyushu, Fukuoka, Japan. [15]SDG LLC, Cambridge, MA, USA. [16]Biogen, Cambridge, MA, USA. [17]AbbVie, Libertyville, IL, USA. [18]Division of Rheumatology, Allergy and Immunology, University of California, San Diego, CA, USA. [19]Janssen Research and Development, Spring House, PA, USA. [20]Department of Dermatology, Hospital of the University of Pennsylvania, Philadelphia, PA, USA. [21]Amgen, Thousand Oaks, CA, USA.

Steering Committee

Alain Cornet[12], Alan Friedman[17], Alessandro Sorrentino[22], Anna Stevens[23], Catherine Barbey[24], Cynthia Aranow[7], Darshini Ayton[2], Ed Vital[8,9], Elaine Karis[21], Eric Morand[1], Erika Noss[19], Eve MD Smith[25], George Stojan[26], Jorge Ross Teres[27], Joy Buie[11], Justine Maller[27], Kate Gregory[1], Kathryn Connelly[1], Kenneth Kalunian[18], Maya Hojnik[28], Nikolay Delev[23], Laura Eades[1], Laurie Burke[4], Maria Dall'Era[29], Patrick Marquis[30], Qing Zuraw[19], Rachel Koelmeyer[1], Rangi Kandane-Rathnayake[1], Richard Furie[31], Ronald van Vollenhoven[32], Sandra Garces[21], Tim Coulom[28], Vera Golder[1], Victoria Werth[20] & Ying Sun[13]

Scientific Advisory Board

Anca Askanase[6], Cailin Sibley[19], Christian Stach[26], Cristina Vazquez-Mateo[13], Elaine Karis[21], Eric Morand[1], Eric Zollars[21], Guillermo Pons-Estel[10], Heath Guay[17], Hermine Brunner[3], Hussein Al-Mossawi[22], Jeanette Andersen[12], Joan Merrill[33], Jorge Ross Terres[27], Justine Maller[27], Karen Costenbader[34], Khadija Dantata[11], Laurent Arnaud[5], Lee Simon[15], Maria Silk[28], Marta Mosca[35], Maya Hojnik[28], Nicki Bush[19], Subhashis Banerjee[23], Thierry Sornasse[17], Tim Coulom[28], Yoshiya Tanaka[14] & Youmna Lahoud[16]

Patient Advisory Panel

Alain Cornet[12], Blanca Rubio[36], Dalila Tremarias[37], Dalilah Kalla[38], Gonzalo Tobar Carrizo[39], Imasha Adisa[11], Jeanette Andersen[12], Joy Buie[11], Khadija Dantata[11], Shiori Nagamori[40], Sibongile Komati[41], Stephanie Scoggins[42], Susanne Udengaard Gydesen[43], Toni Grimes[44], Vinita Haroun[45] & Zoe Karakikla-Mitsa[46]

Industry Advisory Board

Alan Friedman[17], Ann Eldred[17], Bianca Fedriz[23], Catharine Lindholm[22], Catherine Barbey[24], Charlotte Lupton[22], Christian Stach[26], Dennis Grasela[23], Elaine Karis[21], Eric Morand[1], Erik Thomas[13], Erika Noss[19], George Stojan[26], Jorge Ross Terres[27], Justine Maller[27], Maria Juarez[26], Maria Silk[28], Maya Hojnik[28], Nicki Bush[19], Nitin Kumar[23], Oliver Guenther[13], Samantha Pomponi[23], Sanjeev Roy[13], Shelly Kafka[17], Tim Coulom[28], Thomas Morel[26], Youmna Lahoud[16] & Yulia Pincus[19]

Consensus statement

Protocol Working Group

Alan Friedman[17], Anca Askanase[6], Cynthia Aranow[7], Darshini Ayton[2], Ed Vital[8,9], Eric Morand[1], Guillermo Pons-Estel[10], Hermine Brunner[3], Kathryn Connelly[1], Kenneth Kalunian[18], Khadija Dantata[11], Laura Eades[1], Laurent Arnaud[5], Laurie Burke[4], Lee Simon[15], Qing Zuraw[19], Rachel Koelmeyer[1], Sandra Garces[21], Victoria Werth[20], Ying Sun[13], Yoshiya Tanaka[14] & Youmna Lahoud[16]

[22]AstraZeneca, Cambridge, UK. [23]Bristol-Myers Squibb, Princeton, NJ, USA. [24]Biogen, Baar, Switzerland. [25]University of Liverpool, Liverpool, UK. [26]UCB Pharma, Brussels, Belgium. [27]Genentech, San Fransisco, CA, USA. [28]Eli Lilly, Camel, IN, USA. [29]University of California, San Fransisco, CA, USA. [30]Modus Outcomes, Cambridge, MA, USA. [31]Northwell Health, Great Neck, NY, USA. [32]Amsterdam University Medical Centers, Amsterdam, Netherlands. [33]Oklahoma Medical Research Foundation, Oklahoma City, OK, USA. [34]Brigham and Women's Hospital, Boston, MA, USA. [35]University of Pisa, Pisa, Italy. [36]Lupus Europe, Madrid, Spain. [37]Lupus Europe, Cardiff, Wales. [38]Lupus Alert, Beau Bassin-Rose Hill, Mauritius. [39]Agrupación Lupus Chile, Santiago, Chile. [40]Patients Association for Collagen Vascular Diseases Japan (PaCVD Japan) and Nanbyo Support Network Japan, Sapporo, Japan. [41]Lupus Foundation South Africa, Pretoria, South Africa. [42]Lupus Foundation of America, Los Angeles, CA, USA. [43]Lupus Europe, Copenhagen, Denmark. [44]Lupus Foundation of America, Phoenix, AZ, USA. [45]Lupus Canada, Calgary, Canada. [46]Lupus Europe, Leasowe, UK.