Preliminary core sets of measures for disease activity and damage assessment in juvenile systemic lupus erythematosus and juvenile dermatomyositis

Nicolino Ruperto1a, Angelo Ravelli1a, Kevin J. Murray2a, Daniel J. Lovell3b, Boel Andersson-Gare3a, Brian M. Feldman5b, Stella Garay6a, Wietse Kuis7a, Claudia Machado8a, Lauren Pachman9b, Anne-Marie Prieur10a, Lisa G. Rider11b, Earl Silverman5b, Elena Tsitsami12a, Pat Woo13a, Edward H. Giannini3b, Alberto Martini1a for the Paediatric Rheumatology International Trials Organisation (PRINTO)a and the Pediatric Rheumatology Collaborative Study Group (PRCSG)b

Objective. To identify preliminary core sets of outcome variables for disease activity and damage assessment in juvenile systemic lupus erythematosus (JSLE) and juvenile dermatomyositis (JDM).

Methods. Two questionnaire surveys were mailed to 267 physicians from 46 different countries asking each member to select and rank the response variables used when assessing clinical response in patients with JSLE or JDM. Next, 40 paediatric rheumatologists from 34 countries met and, using the nominal group technique, selected the domains to be included in the disease activity and damage core sets for JSLE and JDM.

Results. A total of 41 response variables for JSLE and 37 response variables for JDM were selected and ranked through the questionnaire surveys. In the consensus conference, domains selected for both JSLE and JDM activity or damage core sets included the physician and parent/patient subjective assessments and a global score tool. Domains specific for JSLE activity were the immunological tests and the kidney function parameters. Concerning JDM,
Juvenile systemic lupus erythematosus (JSLE) and juvenile dermatomyositis (JDM) are systemic connective tissue diseases with onset before 18 years of age that are characterized by heterogeneous clinical manifestations, unpredictable courses and substantial risk of morbidity. A rational therapeutic approach for both these diseases is hampered by the fact that information regarding safety and efficacy of drug therapies is mainly based on small, uncontrolled case series or on adult data. One of the main reasons for the difficulties in performing controlled trials is the lack of standardized and validated assessment techniques. This deficiency results in the inability to evaluate therapeutic responses in individual patients and limits the interpretation and comparison of the few trials that have been conducted.

It is now generally agreed that, in order to capture the totality of the effects of a disease or its treatment upon a child or adolescent with a systemic connective tissue disease, the assessment of disease activity, accumulated damage and self-assessed and/or parent proxy-reported quality of life is required [1]. Disease activity has been defined as the ensemble of reversible manifestations resulting directly from the inflammatory disease process, whereas disease damage is reflected in irreversible (or at least persistent) and cumulative change in organ or system anatomy, physiology, pathology or function which may result from prior disease activity, complications of therapy or co-morbid conditions [2]. The health related quality of life (HRQL) is a broad concept that refers to the state of physical, mental, emotional and social well-being and to the family environment [3, 4]. The simultaneous assessment of disease activity and accumulated damage not only allows for a better comparison among therapeutic trials of different agents, but also helps to ascertain whether or not therapies which improve disease activity may result in increased damage. The importance of measuring the HRQL in children and adolescents with a chronic disease has been increasingly recognized in order to understand their perception of the impact of the disease and the therapy on their life. Of note, some of the assessment tools validated for adults may not be suitable for paediatric patients because children and adolescents deserve proper instruments that take into account the disease- and physical/mental age-related issues that are inherent to development.

To address these problems, the Paediatric Rheumatology International Trials Organization (PRINTO) and the Pediatric Rheumatology Collaborative Study Group (PRCSG) undertook a combined effort with the aim to identify, validate and promulgate the core sets of measures for disease outcome assessment to be used in future clinical trials in patients with JSLE and JDM. In this paper, we report the results of the first phase of the project, which was aimed at developing consensus on a list of domains to be included in the preliminary core sets for disease activity and disease damage assessment.

Patients and methods

The project was conducted by using well-recognized consensus formation methodology, specifically designed to combine judgements from a group of experts in a particular field: the Delphi Technique and Nominal Group Technique (NGT) [5, 6]. The Delphi Technique utilizes a series of well-defined questionnaire-based surveys. NGT is a structured face-to-face meeting designed to facilitate reaching consensus on the topic field of study. These techniques have been used to develop the outcome measures of several chronic rheumatic diseases, including juvenile rheumatoid arthritis [7], adult rheumatoid arthritis [8], adult-onset SLE [9, 10] and idiopathic inflammatory myopathies [11]. Consensus formation methodology must be designed so that each step is based on the results of the previous steps.

The project was divided in 2 phases.

Phase 1: postal surveys

Using the Delphi Technique, two sequential questionnaire-based surveys were performed to select and then rank the variables used in routine clinical practice to assess if a patient with JSLE or JDM has responded to a given therapy. The surveys involved 267 experienced practicing paediatric rheumatologists, of whom 178 were members of PRINTO and 89 were members of the PRCSG, from 46 different countries. In the first survey, physicians were asked to indicate up to 10 variables that they judged as clinically most important. Any type of variable (e.g. laboratory tests, questionnaires, indices of disease activity) could be chosen. In the second survey, variables indicated by at least 10 responders in the first questionnaire, and additional variables used in published therapeutic trials or observational studies in JSLE or JDM, were listed in alphabetical order. Among these variables, physicians were asked to select, and rank in order of

Conclusions. We developed preliminary core sets of measures for disease activity and damage assessment in JSLE and JDM. The prospective validation of the core sets is in progress.

Key words: Consensus, Core set, Disease activity assessment, Disease damage assessment, Juvenile dermatomyositis, Juvenile systemic lupus erythematosus.
importance, their top 10 choices. Physicians were also asked to
define the minimum and maximum number of variables that
should be included in each core set and their willingness to use a
dichotomous index (i.e. an index that would classify each
patient as improved or not improved) for the response-to-
treatment evaluation in clinical trials and/or in their routine
clinical practice.

The ultimate goal of the surveys was to obtain information
on the utilization of the variables, i.e. the ability of the variables
included in the core sets to measure disease activity and/or
damage in JSLE and JDM in a sensitive and practical way.
Mailing, e-mail, fax or telephone reminders were used to ensure
a response rate of at least 80% for both surveys.

Phase 2: consensus conference

Following the surveys, a 4-day consensus conference (2 days for
JSLE and 2 days for JDM) was held in Pavia, Italy from March
31 to April 3, 2001. The meeting was attended by 40 experienced
pediatric rheumatologists from 34 different countries,
belonging to either PRINTO (all PRINTO national coordina-
tors) or PRCSG. Three moderators with expertise in NGT
(NR, EHG, KJM) oversaw the implementation of this format
at the meeting. The goal of the meeting was to reach, through
NGT, a consensus upon the domains and variables (see below
for definitions) that should be included in the JSLE and JDM
disease activity and damage core sets. Prior to the meeting
participants received a booklet containing relevant articles
about the potential outcome measures for JSLE [12–20] and
JDM [19–28] and copies of the instruments to be analysed
during the consensus conference. At the meeting, after
introductory lectures concerning the characteristics and scoring
systems of the outcome measures most commonly used in the
assessment of activity and damage in the two diseases under
study, attendees were randomly assigned to three working
groups. Each group worked in a separate room with a
moderator. For every exercise, attendees were first asked to
work individually and then to express their opinion in a guided
discussion. At the end of the work, the results were pooled, and
an 80% consensus from all 40 attendees was required to
consider each problem as solved.

During the meeting, the following five exercises were
accomplished:

NGT exercise 1. Attendees were asked to classify all
variables selected in the two mailing surveys into specific disease
‘domains’. A disease domain was defined as a broad category
comprising more than one variable; some indicative domains
were proposed by the Steering Committee, but attendees were
free to select other domains and/or to modify the definition of
each domain.

NGT exercise 2. Attendees were asked to classify all variables
analysed in the first exercise into the ‘concepts’ of disease
activity and/or damage.

NGT exercise 3. Attendees were asked to select and rank the
domains that should be included in the core sets for disease
activity and/or damage.

NGT exercise 4. Attendees were asked to select the variables
to be used to measure each domain of disease activity and/or
damage core sets.

NGT exercise 5. Attendees were asked to identify the
inclusion and exclusion criteria for the large-scale data
collection study that was planned for the prospective validation
of the proposed core sets.

The booklet with the summary of the consensus conference
is available upon request to PRINTO [29].

Statistical analysis
Analyses of the questionnaire-based surveys and consensus
exercises were descriptive in nature. Weighted averages were
used for the analysis of the second international survey and for
the NGT consensus exercise 3, according to the following
formula: ranks multiplied by the frequency of response divided
by the number of responders.

Results
Phase 1: postal surveys
Response rate in the first survey was 142/178 (80%) among
PRINTO members and 32/89 (36%) among
PRCSG members (174/267, 65% overall). Only the top
variables, chosen by at least 10 responders, were included
in the second mailing survey; these were 41 variables for
JSLE and 37 variables for JDM (data not shown). Response rate in the second survey was 146/178 (82%)
among PRINTO members and 76/99 (77%) among
PRCSG members (222/277, 80% overall). The analysis
of the results according to the responder membership
showed that the same 17 variables were ranked as most
important for either JSLE or JDM by both PRINTO and
PRCSG members (data not shown). The mean number of
variables that should be included in the core set ranged
from 8 to 14 for JSLE and from 7 to 12 for JDM. The
number of physicians who were willing to use a
dichotomous index to classify patients as improved or
not improved was 172/200 (86%) for JSLE and 178/204
(87%) for JDM.

Phase 2: consensus conference
The domains and variables included in proposed core
sets for JSLE disease activity and damage assessment are
reported in Tables 1 and 2, respectively; those for JDM
disease activity and damage assessment are reported in
Tables 3 and 4, respectively. Because the impact of JSLE
and JDM and their treatment in children and adolescents
includes not only the organ damage seen in adult patients
but also the effects on growth and development, a specific
domain ‘growth and development’ was nominated and
agreed by 100% consensus, to be added in the disease
damage core set for both diseases. Consensus on the
variables to measure this domain was reached through
e-mail discussion after the meeting. Selected variables
were: height, weight, menses (regular, irregular or
stopped) and Tanner puberty stage. Furthermore, all
participants agreed that the evaluation of HRQL should
be advised in clinical studies devoted to the assessment of
disease activity and/or damage to capture the influence of
the disease on the patient lifestyle.

Discussion
General comments
Using Delphi and NGT consensus formation techniques,
we developed preliminary core sets of measures for the
assessment of disease activity and accumulated damage
in patients with JSLE and JDM (Tables 1–4). These
proposals are designed to ensure that certain minimum
criteria/standards are applied to future observational and randomized controlled trials and to facilitate comparison or meta-analysis of different studies in the future. Furthermore, they can assist in standardizing outcome measurements in current clinical practice. Each core set includes the indication of instruments that can be used to assess each domain. The specific instruments can be modified or integrated whenever new valid tools or better laboratory indicators are developed either for the measurement of a particular domain or to be more suitable for use in paediatric patients.

The proposed core sets combine aspects of the disease that can be measured easily in clinical practice. Indeed, one of the primary goals of the project was to avoid the inclusion of domains/variables that cannot be assessed reliably world-wide. For example, variables requiring specific imaging techniques (e.g., MRI or bone densitometry) were not included because some technical instruments are not readily available and the use of different instruments to assess the same item (e.g., dual-energy X-ray absorptiometry vs quantitative ultrasound for bone mineral density) may not yield comparable results. The same concerns, however, can be raised for simpler measures such as the anti-DNA antibody level, which can be measured by a variety of different techniques, again non-comparable.
Comments specific to JSLE

The preliminary core set of outcome measures developed for randomized clinical trials and longitudinal observational studies in adult-onset SLE [9, 10] includes four domains: disease activity, health related quality of life, accumulated damage and toxicity/adverse events (including death). A specific domain for toxicity/adverse events was not included in the proposed JSLE core sets because we considered that the pertinent variables are already included in the disease activity or damage global scores. In the prospective phase of the study (see below), the disease activity will be assessed by the Systemic Lupus Erythematosus Disease Activity Index (SLEDAI) [12] (which obtained the highest preference rate in the consensus conference), the European Consensus Lupus Activity Measurement (ECLAM) [15, 16], the Systemic Lupus Activity Measure (SLAM) [14], and the accumulated damage through the Systemic Lupus International Collaborating Clinics/American College of Rheumatology Damage Index (SDI) [18]. All these indices were found to be reliable for use in patients with JSLE [30–33].

A critical issue, however, needs to be addressed as to whether the global scores are really the best way to assess the drug response in a disease such as lupus, which may affect many different organs/systems. Although these scores provide a rough and ready guide to disease activity, they can be of lesser value in trying to capture a variation in disease activity across the organs and systems. To give an example, consider an hypothetical situation in lupus where a new drug is used and found to be helpful in treating skin rash, arthritis and cerebritis, but of no value in treating renal disease: this drug could in fact even be harmful, but the global score system may not capture these nuances and indeed may miss them altogether. To overcome these potential shortcomings, serious consideration should be given to devising more sophisticated tools that not only allow calculation of a global score, but also provide a more detailed and comprehensive assessment and scoring of the disease activity in each organ/system.

In the disease activity core set for lupus, potential areas of controversy relate to the introduction of kidney abnormalities and anti-DNA antibody level. Because many patients with active JSLE do not have kidney abnormalities, it is anticipated that this domain will be applicable only to the subset of patients with renal disease. Furthermore, the role of anti-DNA antibodies as an indicator of disease activity in lupus is rather controversial. Indeed, some studies have found a relatively weak correlation between these antibodies and overall disease activity and their presence has not been consistently shown to predict flares of lupus activity [33]. Nevertheless, due to the foremost importance of anti-DNA antibodies in the assessment of lupus activity in clinical setting, we believe it is essential to include their measurement in the SLE activity core set, expecting to critically evaluate their role and reliability after the prospective data collection for the validation phase. Immune complex-mediated activation of complement through the classic pathway is believed to be one mechanism by which tissue injury occurs in SLE patients. Therefore, serum complement levels are commonly used to determine the disease activity in lupus patients. However, the value of these tests as reliable markers of disease activity and response to therapy has been questioned [34, 35]. Furthermore, hereditary deficiencies in the complement components of the classic pathway have been shown to increase the risk of SLE, indicating that these complement components may exert a protective role against the development of this disease [36, 37]. Antibodies against C1q have been detected in patients with SLE, and some reports have noted a close relationship with renal manifestations [38, 39]. Nevertheless, further studies are needed to establish whether anti-C1q titres are a reliable predictor of the activity of renal disease. The ESR (erythrocyte sedimentation rate) and CRP (C-reactive protein) were not included in the disease activity core set for SLE because they were not considered as reliable indicators of lupus activity. Notably, the CRP is often normal or only slightly elevated in SLE patients and marked elevations of its levels are strongly suggestive of an infection [40].

TABLE 4. JDM disease damage core set

<table>
<thead>
<tr>
<th>Domainsa</th>
<th>Variables</th>
<th>Consensus agreement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global assessment by physicians</td>
<td>Physician’s global assessment of damage (10 cm VAS or Likert scale)</td>
<td>100%</td>
</tr>
<tr>
<td>Functional ability assessment</td>
<td>Childhood Health Assessment Questionnaire (CHAQ)</td>
<td>100%</td>
</tr>
<tr>
<td>Growth and development</td>
<td>Height and weight</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td>Menses (regular, irregular, stopped)</td>
<td>100%</td>
</tr>
<tr>
<td></td>
<td>Tanner puberty stage</td>
<td>100%</td>
</tr>
<tr>
<td>Global JDM damage tool (questionnaire)</td>
<td>Myositis Damage Index (MDI)b</td>
<td>93%</td>
</tr>
<tr>
<td>Muscle strength assessment</td>
<td>Childhood Myositis Assessment Scale (CMAS)</td>
<td>85%</td>
</tr>
</tbody>
</table>

The consensus agreement reported in the third column is the frequency of agreement among the 40 participants in the consensus conference. A minimum of 80% consensus from all 40 attendees was required to consider the problem as solved.

A 98% consensus was reached on the introduction of HRQL assessment in studies on disease damage in JDM.

bThe MDI combines two tools named the Myositis Disease Damage by Visual Analogue Scale (MYODAM-VAS), and the presence or absence of clinical features via a modification of the SDI (with permission from the IMOACSG) [42].
Comments specific to JDM
A preliminary core set of measures for disease activity assessment in the idiopathic inflammatory myopathies has been recently proposed by the International Myositis Outcome Assessment Collaborative Study Group (IMOACSG) [11]. It includes physician and patient/parent global assessments, muscle strength, physical function, laboratory evaluation (muscle enzymes) and an assessment of extra-skeletal muscle disease. Although there are minor differences in the structure of the domains, this activity core set is very similar to ours. This would facilitate the further development of a standardized approach to patients with inflammatory myopathies. In spite of this progress, however, the assessment of JDM patients is still imperfect. To give an example, many measures, including those that evaluate muscle strength and physical function, are indistinguishable between disease activity and prior damage; furthermore, many measures of disease activity are not sensitive enough when accumulated damage interferes with the ability to detect ongoing activity [2]. In our study this uncertainty is reflected by the fact that muscle strength and functional ability assessments were included in both disease activity and damage core sets. Another source of concern is whether the global assessments by physicians, parents or patients are really as reliable as some in the past have claimed them to be. We introduced these measures because we believe it is important to obtain a subjective estimation of the disease impact from both the physician and the parent/patient point of view; this choice was supported by the excellent evaluative properties demonstrated by these measures in previous studies on JDM [22]. Of note however, the parent/patient global assessment was not considered suitable for the damage assessment. Concerning the existing global tools for activity and damage assessment, information is still lacking on their evaluative performance in JDM patients [2]. The prospective data collection that has been started since completion of this study (see below) will provide the data for validation in the paediatric age group. A further criticism of our JDM core sets could be that their assessment requires a significant respondent burden on both the family and the physician. Because it is yet unclear whether some measures are redundant (i.e., manual muscle testing, childhood myositis assessment scale and childhood health assessment questionnaire) or reliable in younger children (i.e., manual muscle Testing), we decided to collect data on all the available JDM measures, expecting to develop more strict response criteria after a comprehensive evaluation of the comparative performance of each instrument in the prospective phase of the study. Several markers of immunologic activation, including soluble interleukin-2 receptor, interleukin-1 receptor antagonist, neopterin, factor VIII-related antigen and lymphocyte subsets, were investigated in JDM and appeared to correlate with overall disease activity [41]. However, since the sensitivity of these parameters is still unknown, they were not considered for inclusion in the JDM activity core set.

Growth and development and quality of life assessments
To better cover the impact of the diseases under study in children and adolescents, the consensus conference attendees agreed to nominate a new domain called ‘growth and development’. This domain was included in the damage core set for both JSLE and JDM and was aimed at assessing growth retardation and pubertal delay. As already mentioned, these are serious consequences of chronic disease in children and may cause problems in physical and psychosocial functioning, particularly for adolescents. It is worth mentioning that the assessment of growth failure and pubertal delay has been included in the Myositis damage index, which is currently in the development process [42, 43].

The importance of measuring the HRQL of children and adolescents with JSLE and JDM to understand the impact of the disease and the prescribed therapies on their life was recognized by all attendees, who agreed that the evaluation of HRQL should be introduced in all clinical studies that include the assessment of disease activity and/or accumulated damage. The Child Health Questionnaire was chosen for the prospective data collection because this is the sole instrument validated in the paediatric age range for which translated versions in all the languages of the study participants are available [20, 44].

Prospective validation
The second phase of the project was started after the consensus conference, with a prospective validation of all the variables included in the core sets through a large-scale data collection from the follow-up of patients with JSLE and JDM by PRINTO and PRCSG members. Data for disease activity will be recorded at baseline and after 6 months in patients meeting the following baseline criteria: (i) diagnosis of JSLE according to the ACR revised criteria [45] or diagnosis of JDM according to the Bohan and Peter criteria [46, 47]; (ii) less than 18 yr of age at enrolment; (iii) active phase of the disease defined by the presence of at least one of the following two criteria: (a) start of corticosteroid therapy and/or new immunosuppressive treatment, (b) major change in the dosage of previous corticosteroid and/or immunosuppressive treatment. The data collection for the damage core sets will be extended through a 1-yr period because a 6-month period is probably too short to allow a reliable estimation of the rate of damage accrual.

After completion of this phase, the performance characteristics (validity, reliability, responsiveness, redundancy) for all variables will be investigated. Next, a second consensus conference will be convened, in which attendees will be asked to examine the ability of each variable to separate patients into improved or not improved in order to develop the final core sets and to reach consensus on the preliminary definitions of improvement for JSLE and JDM.
## Acknowledgements

We are indebted to the members of PRINTO and PRCSG whose enthusiastic responses to the questionnaire-based surveys made this work possible. Surveys were obtained from the following 46 countries: Argentina, Australia, Austria, Belgium, Brazil, Bulgaria, Canada, Chile, Costa Rica, Croatia, Cuba, Czech Republic, Denmark, Finland, France, Georgia, Germany, Greece, Hungary, Israel, Italy, Korea (South), Latvia, Lithuania, Luxembourg, Mexico, The Netherlands, New Zealand, Northern Ireland, Norway, Poland, Portugal, Russia, Saudi Arabia, Singapore, Slovakia, South Africa, Spain, Sweden, Switzerland, Tunisia, Turkey, UK, USA, Yugoslavia. We are also indebted to Dr Luci Gado-West and Dr Anna Tortorelli for their invaluable help in data handling, organization skills and overall management of the project. Supported by a grant from the European Union (contract no. QLG1-CT-2000-00514), by IRCCS Policlinico S. Matteo, Pavia, Italy and by the National Institutes of Health, Bethesda, MD (Grant RO3 AI 44046).

### Key Messages

- PRINTO developed preliminary core sets of measures for disease activity and damage assessment to be used in future clinical trials in patients with juvenile systemic lupus erythematosus and juvenile dermatomyositis.

## References