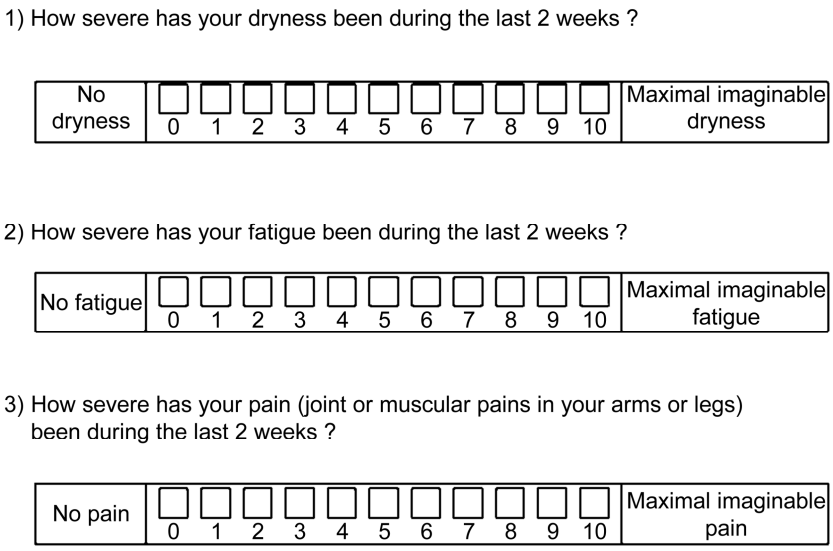

Figure 1 The European League Against Rheumatism (EULAR) Sjögren's Syndrome Patient Reported Index (ESSPRI). The total score is the mean score of the 3 scales.

participating in this international EULAR collaborative project (project code CLI 010).

To be included, patients had to fulfil American-European Consensus Group (AECG) criteria.[9] Additionally, investigators were asked to include approximately half the patients with systemic features. Patients were prospectively followed and two visits were planned at inclusion and at 6 months. No therapeutic intervention was planned in this observational study, and therapeutic management was left to the discretion of the treating physician. This study was conducted with the approval of the institutional review board of GHU Paris Nord (IRB0006477). Depending on local rules, ethical approval has been obtained in other countries whenever necessary. In each country, local ethical requirements have been observed.

Measurement

Disease activity indexes

At enrolment and at 6 months, physicians completed the ESSDAI, the SCAI and SDAI, and assessed systemic disease activity with a 0–10 physician global assessment (PhGA) scale. Also, they evaluated, separately, severity of patients' symptoms with a 0–10 scale (PhGA of patient symptoms).

At 6 months, physicians also had to evaluate the change in disease activity by answering the question 'Compared with the previous visit, is this patient's primary Sjögren's Syndrome activity now…' according to a 5-point Likert scale (much worse, worse, the same, better, much better). Three groups of patients were defined according to change in disease activity: (1) improved, if considered 'better' or 'much better'; (2) stable, if considered 'the same' and (3) worsened, if considered 'worse' or 'much worse.'

Patient-centred measures

At enrolment and at 6 months, all patients completed the ESSPRI (mean score of 0–10 numerical scales for pain, fatigue and dryness features, including oral, ocular and global dryness), SSI, PROFAD questionnaires and a 0–10 patient global assessment (PGA).

At 6 months, patients also had to evaluate the change in their state by answering the question 'Compared to the beginning of the study (6 months ago) how do you evaluate the severity of your Sjögren's syndrome now …' according to a 5-point Likert scale (very importantly improved, importantly improved, slightly improved, no change, worsened). Three groups of patients were defined according to change in symptom state: (1) improved, if considered 'very importantly, importantly and slightly improved'; (2) stable, if considered 'no change' and (3) worsened, if considered 'worsened'.

Objective measures of dryness

At enrolment and at 6 months, objective measures of dryness included Schirmer's dye scores of both eyes, which were considered abnormal if ≤5 mm in 5 min and unstimulated salivary flow (USF), considered abnormal if ≤0.15 mL/min.

Definition of systemic involvement

Systemic involvement was recorded as the presence and/or a past history of the following manifestations: arthritis, myositis, purpura, peripheral or central nervous system, pulmonary or renal involvement, lymphoma or other B-cell proliferative disorder. All these items had been prospectively collected in a standardised case-reported form. Glandular swelling was not considered as a systemic involvement, and was recorded separately.

Statistical analyses

Continuous data are presented as medians with IQR. We used non-parametric tests to analyse continuous variables because the data were not normally distributed.

Correlations between scores

The construct validity was assessed by correlation between the disease-specific indexes and their respective gold standard: the PhGA for systemic disease activity indexes and PGA for patient-centred measures. To assess convergent and divergent validity of disease-specific indices, Spearman's r correlation coefficients were used to assess correlation between all disease activity scores and all patient scores. Higher correlation should need to be observed between scores measuring the same construct (convergent validity), whereas lower correlations might be observed between scores measuring different construct (divergent validity).

Reliability of scores

Reliability was assessed on a subsample of patients with the intraclass correlation coefficient (ICC),[10] as follows: