ABSTRACT

Objectives To describe the development and assess the psychometric properties of the novel ‘Symptoms in Persons At Risk of Rheumatoid Arthritis’ (SPARRA) questionnaire in individuals at risk of rheumatoid arthritis (RA) and to quantify their symptoms.

Methods The questionnaire items were derived from a qualitative study in patients with seropositive arthralgia. The questionnaire was administered to 219 individuals at risk of RA on the basis of symptoms or autoantibody positivity: 74% rheumatoid factor and/or anticitrullinated protein antibodies positive, 26% seronegative. Validity, reliability and responsiveness were assessed. Eighteen first degree relatives (FDR) of patients with RA were used for comparison.

Results Face and content validity were high. The test-retest showed good agreement and reliability (1 week and 6 months). Overall, construct validity was low to moderate, with higher values for concurrent validity, suggesting that some questions reflect symptom content not captured with regular Visual Analogue Scale pain/well-being. Responsiveness was low (small subgroup). Finally, the burden of symptoms in both seronegative and seropositive at risk individuals was high, with pain, stiffness and fatigue being the most common ones with a major impact on daily functioning. The FDR cohort (mostly healthy individuals) showed a lower burden of symptoms; however, the distribution of symptoms was similar.

Conclusions The SPARRA questionnaire has good psychometric properties and can add information to currently available clinical measures in individuals at risk of RA. The studied group had a high burden and impact of symptoms. Future studies should evaluate whether SPARRA data can improve the prediction of RA in at risk individuals.

Key messages

What is already known about this subject?

► A wide range of symptoms can be present in individuals at risk of rheumatoid arthritis (RA), including extra-articular symptoms.

► These symptoms can be severe and disabling.

What does this study add?

► This study used data from qualitative focus interviews to quantify symptoms in individuals at risk of RA.

► The ‘Symptoms in Persons At Risk of Rheumatoid Arthritis’ (SPARRA) questionnaire provides information on location, timing and severity of these symptoms in a large international sample of individuals at risk of developing RA.

How might this impact on clinical practice?

► The SPARRA questionnaire can be used to document symptoms in studies of persons at risk of RA.

INTRODUCTION

A range of symptoms can be present in individuals at risk of rheumatoid arthritis (RA). These individuals are usually defined based on either autoantibody positivity or symptoms. In seropositive at risk persons, symptoms usually occur later than seropositivity. However, information on location, timing and severity of symptoms is still largely lacking.

Symptoms such as joint pain, swelling and morning stiffness represent key elements in the diagnosis of RA. Clinicians have tried to use these and other symptoms to identify those at risk of RA before they fulfill classification criteria for this condition. A European League Against Rheumatism (EULAR) taskforce recently outlined symptoms that were deemed most relevant in differentiating those at risk of developing RA (also known as ‘clinically suspect arthralgia’ (CSA)) from other patients with non-specific joint symptoms. The criteria set for CSA were based...
on expert opinion and have shown value in predicting arthritis.7

Qualitative research in individuals at risk of RA provided a different starting point to evaluate symptoms using the experience of the affected persons to understand the range of their symptomatology. With this approach, multiple focus group interviews were performed in patients with seropositive arthralgia.8–10 Besides symptoms originating from the joints, additional extra-articular themes emerged such as fatigue, distress and loss of motor control, with a reported major impact on daily functioning. The presumed impact of such early symptoms is underscored by increased sick leave and medical ambulatory costs long before diagnosis of RA.11 12

The ‘Symptoms in Persons At Risk of Rheumatoid Arthritis’ (SPARRA) questionnaire was developed based on data from our previous qualitative study.9 The aim of the present study was to describe the developmental process and test the psychometric properties of the SPARRA questionnaire in an international convenience sample of individuals at risk of RA (both autoantibody positive and negative) and to quantify and describe their symptoms based on this questionnaire.

PATIENTS AND METHODS

The development of the SPARRA questionnaire

The content of the SPARRA questionnaire is based on focus group interviews in 15 patients with seropositive arthralgia and 11 patients with early RA with whom initial symptoms prior to the diagnosis of RA were explored.9 These semi-structured interviews were conducted to explore perceptions of symptoms, impact of symptoms and reactions to symptoms and continued until thematic saturation was reached. The content of the questionnaire was also informed by a previous review of the literature related to the earliest symptoms of RA4 and prior research describing domains that were deemed important in predictive algorithms in these at-risk individuals.9 14 The emerging themes were grouped and the most noteworthy and frequently occurring categories were selected. Feedback from the study team (two rheumatologists, one epidemiologist, one expert on psychological testing and two research patient partners) was used to discuss which symptoms to be captured within the questionnaire, discuss realistic timeframes for symptom duration and the number/format of answer categories for severity and impact of the symptoms, taking into account reported variation in each of the domains by individuals from the focus group interviews. Afterwards, four rheumatologists from different countries, working in the field of the at-risk phase of RA but who were not otherwise involved in the project, gave their feedback on the questionnaire. The final questionnaire included 13 symptoms, for which severity and impact were described from none to severe and no to high, respectively. Additional questions were aimed at capturing location and pattern of joint pain (if present) and the presence of morning stiffness. Recently, data on symptoms in the at-risk phase of RA appeared in literature from two other cohorts and one review. These studies contain items on functional limitations, such as difficulty making a fist, which are possibly additive to the SPARRA questionnaire which only contains the following symptoms on function: ‘weakness or loss of motor control’ and ‘impact of symptoms on daily functioning’.15–17 The questionnaire’s design and content was thereafter discussed with patient research partners from both Amsterdam and Birmingham to assess face validity which led to only minor comments and small modifications.

Subsequently, the questionnaire was translated from English into Dutch, Swedish, German and French by at least one native speaker. These native speakers were part of the study teams at the different centres and had knowledge of research in individuals at risk of developing RA, but were not part of the study team that performed the focus interviews leading to the questionnaire development. Thereafter, another researcher from that study team translated it back to English, blinded for the original wording of the items, to complete the formal forward-backward approach as presented by the WHO (steps 1 and 2, except for the fact that back-translation was not performed by a whole expert panel).18 All inconsistencies were resolved in collaboration with a member of the original focus group interview team (LvT), by referring to the original wording in the focus groups. Cross-cultural adaptations were made taking into account cultural aspects of presenting joint symptoms within the different countries.19 As a preliminary pilot test, the Dutch prefinal version of the questionnaire was administered to 30 seropositive individuals with arthralgia from the Netherlands, which did not change the questionnaire (WHO steps 3 and 4, no cognitive interviewing was performed).18 To pre-empt any missing symptoms individuals had the possibility of adding up to two additional symptoms that they thought were relevant. See table 1 for an outline of the questionnaire (complete questionnaire and translated versions presented as online supplementary figures 1–5).

Study participants

To test the psychometric properties of the SPARRA questionnaire, individuals at risk of developing RA defined as individuals with RA-specific autoantibodies (anticitrullinated protein antibodies (ACPA) and/or rheumatoid factor (RF)) or the presence of relevant symptoms (ie, individuals with CSA based on clinical expertise in the different centres with or without RA-specific antibodies) were selected from four European centres: Reade, Amsterdam (n=125) (further called the Netherlands), Sandwell and West Birmingham Hospitals and the University Hospitals Birmingham (n=69) (UK), Karolinska University Hospital (n=15) (Sweden) and the Medical University of Vienna (n=10) (Austria). We have used an international sample of patients at risk for RA, mainly containing consecutive cohort patients (Netherlands, UK, Sweden, in total 88%) and complemented by a convenience sample from Austria and Switzerland. Please note that in Austria, patients were recruited from...
Table 1  Outline of the SPARRA questionnaire

<table>
<thead>
<tr>
<th>Per symptom (see right), the following questions are asked:</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Over the past month how many days of the month have you had (symptom)?</td>
<td></td>
</tr>
<tr>
<td>2. Over the past month how much (symptom) have you had?</td>
<td></td>
</tr>
<tr>
<td>3. What impact has this (symptom) had on your ability to carry out daily activities (eg, work, household chores, childcare, social activities)?</td>
<td></td>
</tr>
<tr>
<td>4. Where did you feel the [symptoms 1 to 9]</td>
<td></td>
</tr>
</tbody>
</table>

**Answer categories:**
- no pain
- mild moderate
- severe
- no impact
- small impact
- moderate impact
- large impact
- hand (one or both), arm (one or both), feet (one or both), leg (one or both)

What is the pattern of symptom development since the time they first began (see patterns on the right; patients could also draw a pattern themselves)

**Note:** full questionnaire (with translations) added as online supplementary material.

**SPARRA, Symptoms in Persons At Risk for Rheumatoid Arthritis.**

their ‘pre-arthritis’ cohort and by additionally searching through their clinical database for patients who were ACPA-positive and/or RF-positive without a diagnosis of arthritis. In Switzerland, individuals could complete a SPARRA questionnaire at any time when they visited for a yearly cohort follow-up. All cohorts were set up to characterise individuals at risk of developing RA and included individuals without prior arthritis (see online supplementary table 1). Arthritis was assessed clinically (by a rheumatologist in all cohorts) by presence of at least one swollen joint: no confirmation by ultrasonography or MRI was used. A cohort of 18 first degree relatives (FDR) of patients with RA from the University Hospital of Geneva (Switzerland) was used as comparison, since this cohort was dealt with separately, since the individuals were recruited based on the fact that they were FDR and not because of symptoms or antibodies and thus included mostly healthy individuals with an increased risk of RA. Individuals were included between November 2014 and December 2016.

**Study procedures**
At baseline, individuals completed the SPARRA questionnaire and had clinical data collected, including antibody status, total painful (tender joint count 44) and swollen joints (swollen joint count 44), family history, symptom duration, smoking status, Visual Analogue Scales (VAS, ranged 0–100) for pain, patient global assessment and fatigue. Detailed data on comorbidities and medication have not been consistently assessed across the cohorts for the present study. A subgroup of individuals had a follow-up measurement (test-retest) after 1 week and 6 months. Questionnaires completed at the time of or after clinical arthritis development (confirmed by a rheumatologist) were discarded for the current analysis. The study was approved by relevant Ethics Committees and all individuals gave written informed consent.

**Psychometric properties**

**Content validity**
Relevant medical articles on symptoms in the at risk phase of RA were compared with the questionnaire items to see if all relevant facets of the construct had been captured.

**Construct validity**
We performed correlations between baseline questionnaire items and clinical parameters that were deemed to be associated based on expert opinion (seven representatives from all centres). We divided these into items with a very close match (concurrent validity, eg, the item fatigue compared with VAS fatigue) and items with less close a match (construct validity). Individuals had to complete the questionnaire within 2 weeks of clinical measurements.
Agreement and reliability

Test-retest analyses were performed in a subgroup of individuals that completed a second test within 7–14 days and/or after 6 months. The retest questionnaire was only sent after return of the first questionnaire. The analyses were performed in questions not containing time elements. Also, we described the scale reliability at baseline, that is, looking at how closely related the items in the questionnaire are as a group.

Responsiveness

This can only be tested in individuals with expected change in their disease status, in our case VAS scores over a time period of 6 months. No formal clinically relevant VAS score changes have been described in individuals at risk of developing RA. We chose a change of 11 mm as sometimes used in an adult rheumatology setting, and we measured a second arbitrary cut-off of 25 mm, since the questionnaire items are on a 4-point scale. We only analysed questionnaire items that were concurrent with the VAS scores. The questionnaire had to be completed within 2 weeks from the clinical measurements.

Statistical analyses

Construct validity

Correlations were calculated using the Spearman’s rank correlation coefficient. Statistical significance was set as a p value less than 0.05. The VAS scores were used as continuous data. Interpretation: 0–0.30 small, 0.3–0.50 medium, 0.50–1 large.23

Test-retest agreement and reliability

The percentage of agreement in the questionnaire items was given for questions on symptom severity and impact, and the Cohen’s weighted kappa (reliability) was measured.24 Interpretation: <0 no, 0–0.20 slight, 0.21–0.40 fair, 0.41–0.60 moderate, 0.61–0.80 substantial and 0.81–1 almost perfect agreement.25 26 Scale reliability was described with the Cronbach’s alpha.

Responsiveness

The Wilcoxon signed-rank test was used to analyse significant differences between questionnaire items over the 6-month period, taking p<0.05 as significance level.

Handling missing data

All questions in the questionnaire follow the same pattern (table 1). If subquestions ‘a’ (duration of the symptom over the past month) were missing and ‘b–d’ were also missing, then ‘a’ was set as 0 days. Equally, if ‘a’ was set as 0 days, then ‘b–d’ were set as none, no impact and not filled in, respectively. Instead, if ‘b–d’ were filled out while ‘a’ was missing, then we assumed the worst case scenario and set ‘a’ as 16–30 days.

All analyses were performed with SPSS V.21 (IBM, Armonk, New York, USA), except for the Cohen’s weighted kappa’s which were computed using Stata V.13.1 (Stata 2013, College Station, Texas, USA).

Frequency and impact of symptoms in the SPARRA questionnaire

Finally, we analysed data from individuals who were ACPA-positive (with or without RF), only RF-positive and those included in the cohort due to specific symptoms (seronegative CSA). Percentages of symptom duration in the last month (dichotomised to 0–15 days and 16 or more days), severity (none/mild vs moderate/severe) and impact (no/small vs moderate/large) were given for these groups and information on joint pain location and patterns was described.

RESULTS

Study population

Two hundred and nineteen individuals completed the SPARRA questionnaire in 4 European centres, with 18 FDR of patients with RA as comparison (online supplementary table 2). Half of the individuals (excluding the FDR) were ACPA-positive (with or without RF), 24% were only RF-positive and 26% were seronegative with CSA (table 2). The mean age of these study participants was 49 years (SD 13.2) and the median duration of symptoms was 20 months (25th–75th percentile 8–56).

Content validity

The items in the questionnaire represent symptoms that are important for individuals at risk of RA.13 14 27 29 Literature search did not identify any additional symptoms describing the at risk phase, except for self-reported functional limitation as part of a tool to detect early inflammatory arthritis30 and difficulty in making a fist at physical examination in individuals with CSA. In addition to the seropositive individuals, we also selected seronegative individuals with CSA and FDRs to achieve a good representation of the at risk population. To make sure that no key symptoms were omitted, in addition to the 13 predefined symptoms, individuals had the option of adding two symptoms. Forty-three out of the 219 individuals used this option (of whom 14 reported two options). Themes (reported at least twice) were: pain/inflammation around tendons or myalgia (n=7), pain only while using the joint (n=4), dry eyes (n=4), functional limitations (n=2), itching skin spots (n=2) and swelling in the groin or legs (n=2). However, many of them did not describe new symptoms (n=16 alternative diagnosis that they felt explained their symptoms such as osteoarthritis or hernia, n=5 explanatory description, n=2 unclear).

Construct validity (relation to clinical parameters)

Analyses were performed in 208 individuals, since 11 did not have clinical data collected within 2 weeks from the baseline questionnaire. Overall, the correlation between questionnaire items and clinical parameters (VAS pain, VAS global assessment and VAS fatigue) was medium to large with Spearman coefficients ranging from 0.38 to 0.63 (online supplementary table 3, all statistically significant). Correlations were higher when strictly looking at concurrent validity (0.58–0.63). The percentage of...
missing values in the questionnaire items was 1% (15 missing questions were set on the worst case scenario, 91 were set on no symptoms) and in the clinical parameters 3%.

**Test-retest reliability and agreement and scale reliability**

Analyses after 1 week were performed in 51 individuals (20 ACPA+, 26 RF+, 5 seronegative) and after 6 months in 90 individuals (37 ACPA+, 30 RF+, 23 seronegative). The median time difference in the latter 90 individuals between the questionnaires was 6 months (25th–75th percentiles: 5–10). Thirty-eight individuals had a retest after both 1 week and 6 months. Overall, the test-retest agreement for the questions was good to excellent (88%–98%) after 1 week, and good reliability with Cohen’s weighted kappa’s between 0.60 and 0.90 was found (table 3). After 6 months, the agreement was 73%–91%, with lower overall kappa’s between 0.09 and 0.62. Less than 1% of the data was missing (14 missing questions were set on the worst case scenario, 30 were set on no symptoms). Subgroup analysis was not feasible due to low numbers. The Cronbach’s alpha for all items on duration, severity and impact was 0.859, 0.874 and 0.908, respectively (0.958 if all items were combined).

**Responsiveness**

Seventy-four individuals had clinical and questionnaire data after a follow-up of 6 months and within 2 weeks apart. Of these, 31 individuals had a VAS pain change of 11 mm and 12 individuals a VAS pain change of 25 mm; equivalent data for VAS fatigue were 32 individuals and 16 individuals (17% missing VAS change scores). The Wilcoxon signed-rank test for the questionnaire items joint pain severity and impact was non-significant when using both cut-off points for VAS pain (p=0.46 and p=0.24, respectively for cut-off 11 mm and both p=0.43 for cut-off 25 mm). Also, no statistically significant differences were found for fatigue severity/impact for the VAS fatigue (p=0.21 and p=0.63, respectively, for cut-off 11 mm, and p=0.11 and p=0.15, respectively, for cut-off 25 mm).

**Frequency and impact of symptoms**

The frequency of symptoms and their impact in the individuals at risk of RA was high (table 4). Overall, presence of symptoms was reported more often by the seronegative individuals with CSA, followed by the ACPA-positive and then the RF-positive group (except for fatigue which occurred more often in ACPA-positive individuals). In all three groups, the percentage of individuals with symptoms at least 16 days in the past month was highest for joint pain (37%–72%), joint stiffness (34%–68%), weakness or loss of strength (21%–35%) and fatigue (28%–39%). The severity and impact were reported similarly across the three groups with the exception of burning
Joint pain was mostly reported in the fingers (ACPA+ 58%, RF+ 52%, seronegative with CSA 65%); however, the percentage of neck and back pain was also high (ACPA+ 39%, RF+ 46%, seronegative with CSA 47% and ACPA+ 50%, RF+ 52%, seronegative with CSA 57%, respectively) (figure 1). Usually, this joint pain was described as aching, symmetric and only one-third reported them as mild. The location of joint pain had a similar distribution across all groups.

Finally, we evaluated the pattern of joint pain in the period preceding the first questionnaire (table 1). Joint pain rapidly increasing and then remaining constant was reported by 9% (pattern A), joint pain gradually increasing over time by 16% (pattern B) and a more intermittent pattern by 53% (respectively 23% and 30% for in between periods without pain (D) and symptoms coming and going but always some pain present (C)) (see table 4 for classification by antibody positivity or negativity). Fourteen individuals (6%) had missing data. The remainder of the individuals chose the option of either drawing a pattern themselves or describing it. Of those, one was similar to A, five similar to C and six to D. Of the remaining 25, 6 had no symptoms, 11 had a peak in the beginning and then declining symptoms (mostly to 0), 3 had a combination of the intermittent patterns, 1 reported an intermittent pattern with no remaining symptoms afterwards, 3 were unclear and the last individual filled in both A and C.
Comparison with FDR of patients with RA

We compared the questionnaire data with 18 FDR of patients with RA. Missing data in this cohort were very low (1 missing question and 13 questions answered with two options were set on the worst case). The prevalence of symptoms was lower, except for numbness and change in skin colour (table 4). The location of joint pain was similar for these individuals (figure 1), just like joint pain patterns A, C and D (11%, 22% and 28%, respectively). The percentage of pattern B was shifted towards the open option where individuals could fill in a pattern themselves (B in 5.6% and other in 33%). Of the latter six individuals, three had no symptoms, two had a peak in the beginning and then no pain and the last can be set as D.

**DISCUSSION**

The SPARRA questionnaire has good psychometric properties. The data show a high burden, severity and impact of symptoms in individuals at risk of developing RA.

Evaluation of a complete set of symptoms that might be predictive for RA development is a challenge in the setting of prospective cohort studies, since questions need to be predetermined and are usually based on classification criteria. Using the SPARRA items derived directly from ACPA-positive symptomatic at-risk individuals gave the opportunity to gain more insight into these symptoms. Questionnaires including symptoms have not yet been reported in cohorts researching individuals at risk of RA. Some use generic tools addressing functional limitations rather than specific symptoms, for example, the 36-item Short Form survey and EuroQol five dimensions questionnaire, which have shown their relevance in patients with RA and other musculoskeletal disease, but not yet in the at risk phase.31–33 Recently, the use of the Health Assessment Questionnaire (HAQ) has been described in individuals with CSA and it was shown that overall a low score (median 0.5) was present, but a higher score seemed to correlate with inflammation on MRI and arthritis development.17
Construct validity and responsiveness of the SPARRA questionnaire were moderate and non-significant (respectively) in this study. The moderate correlation between clinically used VAS scores with questionnaire items can partly be explained by the fact that VAS scores measure symptoms from the past week compared with SPARRA questionnaire items measuring symptoms during the last month. Alternatively, the questions may measure different elements and reflect symptom content not captured with regular VAS pain/global assessment. This would mean that the questionnaire adds information to currently available clinical measures in individuals at risk of RA. The low responsiveness in individuals with changing VAS pain and fatigue scores might relate to the fact that no formal cut-offs are described for this study population and responsiveness could only be measured in a small subgroup. It could be that a change in the questionnaire items is not useful in follow-up of individuals at risk of RA; however, future evaluation in a larger population in a longitudinal setting is necessary.

The data showed that besides the expected joint symptoms, the burden of general and nervous system-related symptoms such as burning and tingling sensations, numbness and fatigue was high, especially among the ACPA-positive individuals. An explanation could be the presence of a more general subclinical inflammation as was suggested by MRI and PET studies as well as an early involvement, prior to subclinical inflammation, of the nervous system in ACPA-positive individuals as both in-vivo and in-vitro studies have suggested. Higher scores in joint pain and joint stiffness in the seronegative individuals might be a consequence of the fact that these individuals were included mainly based on the presence of symptoms. This was underscored by a recent cohort study in which a difference was shown between the symptomatic phase preceding ACPA-positive and ACPA-negative RA, and seronegative individuals had more symptoms at baseline. A lower burden of symptoms in the cohort of FDR of patients with RA was expected as these individuals mostly were without symptoms or antibodies at completion of the baseline SPARRA questionnaire. A similar joint pain location and pattern in the FDR’s as compared with the seropositive arthralgia group may reflect their increased risk for RA.

For missing data imputation in this study, we decided to use a worst case scenario approach. We also checked whether using a best case scenario changed the results, which was not the case (data not shown).

A limitation of the study may be selection bias caused by partly using a convenience sample in which non-response and the associated reasons are lacking. This may have led to overestimation of the burden of disease, since individuals with more symptoms may be more willing to complete the questionnaire. However, since the (heterogeneous) study population was taken from a set of individuals found in daily practice in secondary care, we expect that the results from the study can be generalised to other secondary care practices. Also, 88% of individuals were assessed in a consecutive manner. Another limitation is the fact that comorbidities may influence the reported symptoms and data on comorbidities were not collected for the present study.

The predictive value of the questionnaire items (possibly combined with items of the HAQ) for developing RA requires further investigation. This can hopefully guide item reduction of the questionnaire in the future. Furthermore, responsiveness could also be investigated in medication trials for preventing RA and a longitudinal design would be helpful to assess its use.

In conclusion, this study provides evidence of good psychometric properties of the SPARRA questionnaire, except for moderate construct validity and low responsiveness. In individuals at risk of RA, symptoms are frequent and severe and have a high impact. Future studies are needed to evaluate whether data from the SPARRA questionnaire can help to improve the prediction of RA.


Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977;33:159–74.


