Rheumatoid Arthritis Severity Scale: a brief, physician-completed scale not confounded by patient self-report of psychological functioning

W. A. Bardwell¹,², P. M. Nicassio¹,², M. H. Weisman³, R. Gevirtz² and D. Bazzo⁴

¹Department of Psychiatry, University of California, San Diego, ²California School of Professional Psychology, San Diego, ³Department of Medicine, University of California, San Diego and ⁴Department of Family and Community Medicine, University of California, San Diego, California, USA

Abstract

Objective. The purpose of this study was to develop a brief measure of severity for rheumatoid arthritis (RA) that would not be seriously confounded by psychological functioning. The Rheumatoid Arthritis Severity Scale (RASS), designed for use by physicians on their own patients, consists of three visual analogue scales: Disease Activity, Functional Impairment and Physical Damage.

Methods. Ninety-four RA outpatients completed the Health Assessment Questionnaire (HAQ) Disability, Pain Severity, Health State subscales and the Symptom Checklist-90—Revised (SCL-90-R) Anxiety, Depression and Somatization subscales. Rheumatologists completed the RASS on their own patients.

Results. Results suggest that the RASS is internally consistent (α = 0.85) and valid. RASS Disease Activity, Functional Impairment, Physical Damage correlated with HAQ Disability (r = 0.40, 0.68, 0.61; P < 0.01), Pain (r = 0.37, 0.34, 0.34; P < 0.01) and Health State (r = −0.27, −0.36, −0.27; P < 0.01). RASS Physical Damage uniquely predicted longer illness duration (years with RA). In contrast to the HAQ, RASS subscales shared less variance with anxiety, somatization and depression scores.

Conclusions. Preliminary data suggest that the RASS may be a quick, reliable, valid physician-completed RA severity scale that compares favourably with the longer, patient-completed HAQ.

KEY WORDS: Rheumatoid arthritis, Disease severity, Health status.

Rheumatoid arthritis (RA) has a chronic and progressive course with a profound psychosocial and economic impact [1]. The assessment of its severity is extremely important for monitoring the clinical course of the disease, gauging the effectiveness of medical, pharmaceutical and behavioural interventions, and quantifying the impact of pathophysiological and biopsychosocial correlates.

Wolfe [2] proposed that severity is equivalent to RA outcome, consisting of objective components (remission, physical damage, acute-phase reactants, joint swelling) and objective plus subjective components (joint tenderness, pain, grip strength, fatigue, functional/work disability, global severity, adverse drug reactions, costs, mortality, social effects). Fries et al. [3] reported five components of RA outcome: disability, pain, cost, iatrogenic reactions and mortality. In the present study we considered three of the more commonly mentioned severity dimensions: disease activity, functional impairment and physical damage.

RA disease activity can be defined as the potentially reversible manifestations of inflammation: pain, stiffness, fatigue, joint swelling, weight loss, elevated erythrocyte sedimentation rates (ESR) and anaemia. During the course of RA, disease activity may fluctuate widely. Existing measures to assess disease activity include joint scintigraphy, magnetic resonance imaging (MRI), counting tender/swollen joints, blood tests for acute-phase reactants, and the quantification of morning stiffness or pain [4–8]. Functional impairment refers to diminished ability to perform activities of daily living, employment and other tasks. Most RA patients show some functional impairment not explained by age alone [9, 10]. Measures of physical damage reflect the irreversible physiological effects
of RA itself or iatrogenic sequelae of treatment, by assessing cumulative scarring, destruction and deformity of joints and underlying bone. These are best assessed radiographically [11–13].

The most complete method of measuring RA severity is based on American College of Rheumatology (ACR) standards [14] and involves clinical assessment (history, physical examination), laboratory tests (e.g. ESR) and imaging procedures (e.g. X-rays, MRI). While accurate, such thorough assessments are time-consuming and costly. The literature suggests that portions of these detailed assessment methods can be used independently. For example, a recent study found that clinical examination of the hands and feet correlated significantly ($r = 0.79$ and 0.66 respectively) with the X-ray Larsen score [15]. ESR has also been used by itself as an objective measure of RA activity [16–18] and has been shown to correlate with muscle strength ($r = 0.37$) [16] and with the Rapid Assessment of Disease Activity in Rheumatology (RADAR) scale ($r = 0.35$), a self-report measure [19].

While accuracy is important in gauging disease severity, limited resources may not always permit detailed assessments. Also, the incremental accuracy that comes with a more thorough assessment may not always be warranted. Various physician-completed pencil-and-paper measures are valid and reliable alternatives. For example, Rider et al. [20] assessed the validity and reliability of physician-completed global assessments of disease activity and physical damage using visual analogue scales (VAS) for the juvenile idiopathic inflammatory myopathies. They found both measures to be valuable for quantifying therapeutic response.

Patients are often asked to self-rate their disease activity or functional disability using pencil-and-paper scales. These include the Health Assessment Questionnaire (HAQ) [3], the Arthritis Impact Measurement Scale [21], RADAR [19] and the RA Disease Activity Index [18]. However, many pencil-and-paper measures are lengthy and potentially confounded by reporting biases that result from psychosocial factors. For example, some patients may over-report symptoms or disability due to secondary gain or a tendency to somatize [22]. Others may under-report symptoms if they have not accepted their condition or to please their physician—a ‘socially desirable’ response style [23]. In RA specifically, the impact of psychosocial factors on symptom reporting has been well documented [24–27]. Although physician-completed instruments may be influenced less by a patient’s psychological profile, physicians may not have time to complete detailed instruments.

These limitations have created a need for a brief measure of RA severity that is not seriously confounded by the patient’s psychological functioning. The purpose of this study was to devise an efficient, physician-completed measure of RA severity. While still subjective, physician-completed measures are less likely than patient-completed measures to be influenced by the patient’s emotional state or personality patterns. However, in many settings, existing physician-completed instruments are too lengthy and time-consuming. The Rheumatoid Arthritis Severity Scale (RASS) (Fig. 1) assesses disease activity, functional impairment and physical damage. We examined the psychometric properties of this measure by comparing it with existing measures of RA severity, pain, functional disability, general health state and psychological variables. Our findings suggest that, with further validation, the RASS may be an efficient and effective tool for both clinical and research purposes.

Patients and methods

Patients

Patients were recruited consecutively from a university-centred private practice of academic rheumatologists. To qualify, volunteers had to be 18 yr or older, with an established RA diagnosis and under the active care of a rheumatologist. Billing records and medical charts were reviewed by the patient’s rheumatologist to validate the diagnosis. All participants fulfilled ACR criteria for RA. Patients received a consent form outlining their right to refuse to participate in or withdraw at any time from the study.

Of the 120 patients we approached, 85% agreed to enrol in the study, reflecting the willingness of this population to participate in research. Of the 102 participants enrolled, 94 (92%) provided all the required information. Table 1 shows descriptive statistics for demographic variables for the 94 subjects completing the study. Reflecting the RA population as a whole, participants were predominantly middle-aged Caucasian females who have been living with RA for an average of 14 yr.

Procedures

Patients who agreed to participate were mailed a questionnaire package which included a request for demographic information plus the following three self-report instruments.

Symptom Checklist-90—Revised (SCL-90-R) Anxiety, Somatization and Depression subscales. The SCL-90-R consists of 90 physical and psychological symptoms. Each item is rated by the patient on a scale of 0–4. It has been shown to be a reliable and valid measure of anxiety, depression and somatization [28, 29].

HAQ Disability, Pain Severity and Health Status subscales. The HAQ Disability subscale consists of 20 questions in eight functional categories, which patients rate on a scale of 0–3. It has been shown to be a reliable (0.85–0.99) and valid (0.71–0.95) measure of RA functional disability and overall severity. The HAQ Pain Severity and Health Status subscales are both rated using a VAS [3, 30]. To obtain a general index of health status, we standardized and summed the three HAQ subscales (after reversing scores for Health State) into HAQ Total.
### Rheumatoid Arthritis Severity Scale

Use the following scales to indicate the level of Rheumatoid Arthritis severity this patient has experienced in the last month. (If you haven’t had contact with the patient in the last month, then consider their severity as of their last appointment with you.)

The literature shows three dimensions of severity: disease activity, functional impairment, and physical damage. For each, put a slash through the line to indicate this patient’s level of RA severity compared with that of all other RA patients you have seen.

#### Disease Activity:

Taking into account joint counts and sedimentation rates, indicate your assessment of disease activity experienced by this patient.

<table>
<thead>
<tr>
<th>No Disease Activity</th>
<th>Worst RA Disease Activity I have ever seen</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### Functional Impairment:

Taking into account your assessment of the patient’s capacity for self-care, household tasks, social activity, work and physical dexterity, indicate the level of functional impairment experienced by this patient.

<table>
<thead>
<tr>
<th>No Functional Impairment</th>
<th>Worst RA Functional Impairment I have ever seen</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### Physical Damage:

Taking into account your assessment of x-ray criteria (such as erosions), indicate the level of anatomic damage to the joints experienced by this patient.

<table>
<thead>
<tr>
<th>No Physical Damage</th>
<th>Worst RA Physical Damage I have ever seen</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

© 1998 Wayne A. Bardwell, Ph.D.

Fig. 1. RASS. The lines should be 10 cm long.

The RASS. Three board-certified faculty rheumatologists (all in practice locally for ≥25 yr) completed the RASS for each of their patients in the study based on their own prior clinical knowledge of the patient and a chart review, if necessary, including laboratory and X-ray data. Data collection was consistent with usual practice standards. The rheumatologists considered their relationships with most of these patients...
Table 1. Demographic variables

<table>
<thead>
<tr>
<th></th>
<th>Range or no.</th>
<th>Mean</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (yr)</td>
<td>21–83</td>
<td>53.7</td>
<td>14.6</td>
</tr>
<tr>
<td>Years with RA</td>
<td>0.2–49</td>
<td>13.9</td>
<td>11.7</td>
</tr>
<tr>
<td>Number of rheumatology doctors’ office visits (annualized)</td>
<td>0–22</td>
<td>4.2</td>
<td>3.4</td>
</tr>
<tr>
<td>Females</td>
<td>78 (83%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Males</td>
<td>16 (17%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ethnicity, n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caucasian</td>
<td>73 (78%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td>8 (9%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>African-American</td>
<td>2 (2%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>11 (11%)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

to be strong and durable, having treated them for 1–20 yr (mean 9 yr), averaging four to five office visits per year, and having seen the patients, on average, within 45 days of this rating. The RASS consisted of three VAS using 10-cm lines representing disease activity, functional impairment and physical damage (Fig. 1). Scale end-points were described as follows: 0 = no disease activity (functional impairment or physical damage); 100 = worst RA disease activity (functional impairment or physical damage) I have ever seen. The subscales were summed to yield an index of overall health status: RASS Total. Rheumatologists were given general instructions (Fig. 1).

Because many patients did not have current X-rays, and X-rays were not taken for the study, each rheumatologist was asked to estimate the degree of physical damage. Where X-rays were available, they were used in making the ratings.

Data were analysed with SPSS 7.0 (release 7.0; SPSS, Chicago, IL, USA) using correlation, reliability, and multiple regression analysis. Pearson product moment coefficients were used throughout. The validity of the RASS subscales and RASS Total was evaluated by examining their associations with the HAQ Disability, Pain, Health State and HAQ Total; the SCL-90-R Depression, Anxiety and Somatization subscales; and Years with RA. We postulated that, as a measure of disease severity, the RASS would be correlated with HAQ Pain, Disability, Health State and HAQ Total, and with number of years since RA diagnosis. We did not hypothesize significant associations between the RASS and psychological indices.

This was a convenience sample from a clinical population taken over a specified time period. We took all participants who met ACR diagnostic criteria and who agreed to participate, resulting in n = 94. The recommended sample size (assuming medium effect size) is 85 for bivariate correlations and 91 for regression with five predictors. Therefore, this study had sufficient power to detect clinically meaningful effects [31].

Results

Table 2 shows descriptive statistics for the HAQ, RASS and SCL-90-R subscales. The patients in our sample covered nearly the full possible range of scores for each subscale. This indicates that, in our sample, patients ranged from having no or little impairment or symptoms to showing severe symptoms or the maximum level of impairment or symptoms.

Reliability

Table 3 shows correlations among the RASS subscales. The significant interscale relationships constitute evidence that the subscales are measuring the same underlying construct. Functional impairment and physical damage showed the strongest relationship (r = 0.84, P < 0.01), disease activity and functional impairment the next strongest (r = 0.66, P < 0.01) and disease activity and physical damage showed the lowest correlation (r = 0.48, P < 0.01). RASS Total, the sum of the three RASS subscales, was used to assess internal consistency reliability, i.e. the degree to which the subscales measure the same underlying construct (coefficient α). A coefficient α of 0.85 was achieved, reflecting a very high level of internal consistency. By comparison, HAQ Total yielded a coefficient α of 0.70.

Validity

HAQ. Because the HAQ is recognized as a reliable and valid measure of RA health status, we examined correlations between the RASS and HAQ. As evidence of validity, RASS subscales were expected to be positively related to HAQ subscales. Table 4 shows significant relationships between all RASS subscales and the HAQ. The strongest associations were between the RASS and HAQ Disability; weaker relationships were between the RASS and HAQ Health State (all P < 0.01). The highest correlation was for RASS Functional Impairment and HAQ Disability (r = 0.68), an expected result because the subscales measure the same variable (Fig. 2). Also, RASS Total was highly correlated (P < 0.001) with HAQ Total, but was not as strongly related to HAQ subscales as the individual RASS dimensions. In summary, higher scores on RASS Disease Activity, Functional Impairment, Physical Damage and Total were related to higher HAQ Pain, Disability and HAQ Total and lower HAQ Health State.

Years with RA. Because RA usually has a progressive course, a measure of RA severity should be correlated positively with the number of years since diagnosis. Because the rate of functional decline is affected by numerous factors, this relationship should be moderate, not high. Table 5 shows zero-order correlations among the RASS, HAQ and Years with RA. RASS Physical Damage, RASS Functional Impairment and HAQ Disability were correlated with longer times since diagnosis. In contrast, RASS Disease Activity and HAQ Pain and Health State were uncorrelated with Years with RA.

To determine the unique predictors of Years with RA, the RASS and HAQ subscales that showed significant zero-order correlations with Years with RA competed for entry in a stepwise regression model. RASS Physical
Damage ($b = 0.45, t = 4.82, R^2 = 0.20, P < 0.001$) was the only variable to enter the model. Thus, in the presence of the RASS, the HAQ subscales had no independent relationship with Years with RA.

**Psychological factors.** Although chronically ill patients often experience comorbid dysphoric mood and/or a tendency towards somatization, a valid measure of RA severity should not be confounded with generalized psychological distress. Correlations between the RASS, HAQ and SCL-90-R subscales appear in Table 5. Compared with the RASS, HAQ subscales show greater overlap with SCL-90-R subscales. Whereas none of the RASS subscales correlated with Anxiety or Somatization, HAQ Health State correlated negatively with Anxiety and Somatization, and HAQ Disability and Pain correlated positively with Somatization. In contrast, all three HAQ subscales, and RASS Functional Impairment and Physical Damage, correlated positively with SCL-90-R Depression. However, when HAQ and RASS subscales competed for entry in a stepwise regression model predicting SCL-90-R Depression, only HAQ Health State ($b = 20.45, t = 24.77, P < 0.01$) uniquely contributed to depression scores ($R^2 = 0.20, F = 22.80, P < 0.01$). Thus, the HAQ proved more sensitive to depressed mood than the RASS.

**Discussion**

This study presents preliminary data on a new measure of RA severity that shows promise for use in clinical and research settings in which physicians rate RA severity in their own patients. Clinically significant data on the reliability of the RASS are presented. The three RASS subscales were positively correlated with each other, displaying evidence that they tap a common underlying construct. Yet, the relationships were not so strong as to indicate redundancy and to prevent interpretations of individual RASS dimensions. Furthermore, the internal consistency of the RASS was quite strong, and surprisingly higher than that of the HAQ.
The construct validity of the RASS was evaluated by examining its relationship with the HAQ. All RASS subscales and RASS Total were correlated in the expected direction with HAQ Pain, Disability, Health State and HAQ Total. Greater RA severity was associated in all instances with lower self-reported health status. This finding is particularly noteworthy, given the widespread acceptance of the HAQ as a measure of general health status. Moreover, because the correlations emerged across self-report and physician-rated modes of assessment, common method variance cannot explain these results. Although the RASS and HAQ are not parallel measures, the substantial correlation between RASS Functional Impairment and HAQ Disability and the lower correlations among the other scales measuring different constructs reflect a pattern that is suggestive of the convergent and discriminant validity of the RASS.

The significant relationship between RASS Physical Damage and Years with RA is additional evidence of the scale’s construct validity. A substantial amount of evidence has shown that RA duration is associated with progressive deterioration in physical damage to the joints and related comorbidities [10, 32, 33]. The results of this study showed that RASS Physical Damage was a unique predictor of Years with RA when it competed with the HAQ in a stepwise regression model. Prospective research in which the RASS evaluates changes in RA severity over time is warranted to further document its validity. This should include ongoing comparisons against standard outcome measures, such as Disease Activity Scores and ACR response criteria, and radiographic information.

As evidence of discriminant validity, the RASS showed minimal relationship with indices of psychological distress. In contrast to the HAQ, the RASS was uncorrelated with SCL-90-R Anxiety and Somatization subscales. Scores on the RASS are thus free of biases resulting from anxious affect and the tendency to report physical symptoms. This is the pattern that is desired of a measure that purportedly assesses disease activity, and should be reassuring to physicians who would use the RASS to focus their judgements on the health status of RA patients.

It should be noted, however, that all three HAQ subscales and two RASS subscales (Functional Impairment, Physical Damage) did correlate with the SCL-90-R Depression subscale. A substantial amount of previous research has shown positive relationships between disease indices, disability and depressed mood [34–36]. On the basis of this body of evidence, the RASS would be expected to correlate with depression. However, regression analyses confirmed that when the HAQ and RASS competed to predict depression, only HAQ Health State emerged as a unique contributor. It is possible that a longer measure, such as the HAQ, may be a more sensitive indicator of depression than the RASS. Alternatively, the stronger relationship between the HAQ and depression could reflect, in part, a common reliance on self-report. Interestingly, the RASS Disease Activity subscale appears to be totally mood-free in that it did not show significant correlations with any of the SCL-90-R measures.

The value of the RASS as a unidimensional scale was assessed by combining its subscales into a composite score (RASS Total). As expected, this composite score showed a strong, positive correlation with another index of general health status applicable to rheumatic disorders, the HAQ Total, but was less correlated with individual HAQ subscales. These results suggest that unique information is provided by the individual subscales, and some information about specific aspects of functioning is lost, therefore, by the adoption of the summary index. Thus, the high internal consistency of the RASS and its correlation with HAQ Total support the use of the RASS Total as a general index of health status, but not as a surrogate for the individual subscales in predicting individual components of RA severity.

Limitations

Disease activity was assessed by the physicians in retrospect (relying on chart review and knowledge of the patient), a potential concern about the reliability of the data. Whereas the HAQ was mailed to participants upon enrolment into the study, the RASS was completed on the basis of the last visit to the physician. The last visit was an average of 45 days prior to the physician completing the RASS. However, these were generally stable patients and our assessment is that this is not a major weakness. In the future, experimental error could be reduced even further, however, by having the RASS completed immediately after the patient’s visit, at which time the patient would also have completed the HAQ or other validation instruments. It is expected that concurrent administration of the HAQ and RASS in the present study would have strengthened the relationship between the two measures, as time would have been held constant. The fact that significant relationships were derived even after such a time lapse can be construed as additional support for the validity of the RASS.

A potential new use of the RASS would be for it to be completed by physicians who are not familiar with the patients being rated, on the basis of chart review alone. While not within the scope of the current study, this is an important factor to be considered in future studies of this instrument and should be tested prior to generalizing the findings from this study.

For most participants, X-ray evidence was available in the chart and used in the RASS assessment. However, for some participants, the physicians had to rely on other aspects of their clinical judgement, including their knowledge of the patient, to assess physical damage. Future validation studies would be well-served by having X-rays available for all participants. In addition, it would be helpful to have further physiological data (e.g. ESRs) available to further validate the RASS Physical Damage and Disease Activity subscales. Additional reliability data, e.g. intra- and inter-rater reliability, would provide important additional support for this instrument. Also, in composite measures, such as...
the RASS and HAQ, the individual elements may measure different things at different times. Therefore, any correlations must be understood in that context.

Conclusions
These findings have illustrated the potential utility of a brief, physician-completed measure of RA severity applicable to their own patients. The RASS showed superior internal consistency and increased sensitivity to RA duration compared with the HAQ and demonstrated less overlap with psychological constructs. While important evidence for reliability and validity has been provided, further research evaluating the psychometric properties of the RASS should be pursued. For example, although the RASS has a high level of internal consistency, analyses of the scale’s inter-rater and test-retest reliabilities should be conducted. Positive results would suggest that the RASS could be used by physicians to rate RA severity in patients other than their own (e.g. based on chart review). In addition, the use of RASS by non-physicians equally familiar with their own patients (e.g. nurse practitioners) may be an important area for further study. Furthermore, studies examining the correlation of the RASS subscales with functional measures (e.g. grip strength, walking time) and biological indices (e.g. ESR) that are directly reflective of disease processes would add significantly to our understanding of the validity and effectiveness of the RASS as a clinical and research instrument.

The RASS is a clinical tool developed entirely from the judgements of rheumatologists caring for their own patients. It avoids some of the limitations of existing self-report measures and appears quite adaptable for either clinical or research use by physicians in clinical settings with their own patients. While the RASS has shown good correlations with validated patient self-report measures of RA severity, whether it will serve as a substitute for outcome measures in clinical trials will be determined when it is evaluated against standardized outcome measures (e.g. ACR response criteria). While its capacity for documenting clinical change in prospective studies would make the RASS a valuable research tool, we recommend that the RASS be used as a clinical index until additional data are available. We hope these findings will spur others to join us in the further evaluation of this instrument.

References