Are Patient Questionnaire Scores as “Scientific” as Laboratory Tests for Rheumatology Clinical Care?

Theodore Pincus, M.D.

Abstract

Modern medical care is based largely on laboratory advances, such as microbiological cultures giving rise to antibiotics and hemoglobin A1c leading to “tight control” of diabetes, among many others. Development of a “gold standard” laboratory test has appeared attractive for care of patients with rheumatoid arthritis (RA) since rheumatoid factor was identified in the 1940s. Indeed, rheumatoid factor, anti-cyclic citrullinated peptide (anti-CCP) antibodies, erythrocyte sedimentation rate (ESR), and C-reactive protein (CRP) are abnormal in most RA patients. However, each of these tests is normal in at least 30% of patients, and no laboratory test (or any other measure) can serve as a single “gold standard” measure for all individual RA patients. A new approach to quantitative assessment in rheumatic diseases involves patient self-report questionnaires as standardized, quantitative, cost-effective “scientific” data from a medical history, the primary source of RA management decisions. Patient questionnaires distinguish active from control treatments in RA clinical trials at levels similar to laboratory tests (or formal joint counts), and are far more significant in the prognosis of work disability, costs, and premature death than laboratory tests or radiographic scores. RAPID3 (routine assessment of patient index data) on an MDHAQ (multidimensional health assessment questionnaire) requires 5 seconds to score, compared to 114 seconds for a DAS28 (disease activity score). Patient questionnaires do not replace further medical history, physical examination, laboratory or other tests, and require physician interpretation for patient management, as do laboratory tests and all quantitative data. Advances in therapy require laboratory science, but patient questionnaires provide optimal “scientific” data for clinical care.

Modern medical care over the last 150 years has been advanced by the “scientific method,” predominantly based on laboratory tests ranging from microbiological cultures to hemoglobin A1c and many others. Diabetes was recognized in ancient cultures on the basis of a sweet taste of urine (Table 1). The sweet taste was documented, in 1776, to be due to “an excess of a kind of sugar in blood and urine,” and the increased glucose content of diabetic blood was demonstrated by Claude Bernard in 1859. Precise measurement of glucose led to application of the discovery of insulin by Banting and Best in 1922, for which Banting was awarded the Nobel Prize in Medicine in 1923. Hemoglobin A1c was described, in 1958, and elevated levels of hemoglobin A1c in people with diabetes reported in 1969. Hemoglobin A1c was documented to provide a superior method for monitoring control of glucose metabolism in 1976.

Patients with diabetes are now monitored with the goal of “tight control” of abnormal laboratory tests, often primarily through self-management, sometimes with high-technology pumps, with substantially improved outcomes. An approach similar to diabetes has been applied to a strategy of “tight control” of other chronic diseases, including hypertension and hyperlipidemia, resulting in important advances in care and outcomes.

The model of diabetes has appeared attractive to apply the scientific method to diagnosis and management of rheumatoid arthritis (RA). Arthritis also was recognized in ancient cultures, although, rather than RA, it may likely have been gout, ankylosing spondylitis, or osteoarthritis. RA was described by Garrod in 1890. Rheumatoid factor was described, in 1940, by Waaler in Norway and, in 1948, by Rose, Ragan,
Inclusion of patient self-report information, in addition to laboratory tests and joint counts, to assess and monitor patients with RA presents a major paradigm shift from the “biomedical model,” the dominant paradigm of 20th century medicine. This model suggests that information from a patient history is “subjective” and nonquantitative, in contrast to scientific “objective” data from laboratory tests, imaging, and other high-technology sources. Patient information has been regarded primarily, if not entirely, as “clues” to identify appropriate objective “scientific” high technology laboratory and imaging tests for diagnosis and treatment. However, the RA Core Data Set regards patient questionnaires as standardized, quantitative “scientific” data for medical care.

Patient questionnaire scores as “scientific” data, according to the scientific method, are based on the science of psychometrics, developed by social scientists over the 20th century. Psychometric approaches use standardized methodology to collect quantitative patient self-report data at baseline and for comparison from one visit to the next, just as laboratory tests. Indeed, standards of measurement accuracy, validity, reliability (reproducibility), and precision applied to patient questionnaires generally are considerably more stringent than those applied to laboratory tests, possibly, in part, due to the “subjective” source of the data.

A major milestone in development of patient questionnaires for rheumatoid arthritis was publication of the health assessment questionnaire (HAQ) in 1983.34 The Arthritis Impact Measurement Scales (AIMS) in the same issue of Arthritis & Rheumatism in 1980. Shortly thereafter, a few (but not many) rheumatologists, including the author and Frederick Wolfe, adopted these questionnaires for usual clinical care. The HAQ was more easily completed by patients in usual care settings, although the AIMS had superior psychometric properties.36 The author developed a modified version of the HAQ (MHAQ) with eight rather than 20 activities, to allow additional information within a two-page format, reported in 1983.34 Since page 131 Bulletin of the NYU Hospital for Joint Diseases 2010;68(2):130-9, patient questionnaires have been viewed by the rheumatology community (and continue to be viewed by many
rheumatologists) as a poor surrogate for traditional “objective” laboratory tests, formal joint counts, a radiograph, etc. In 1984, it was reported that patient questionnaire scores indicating poor functional status predicted survival of less than 50% of RA patients over the next 5 years (Fig. 1A),37 which had not been reported previously (or since) for baseline radiographic scores or laboratory tests (Fig. 1B indicated significantly higher survival according to more years of formal education, which is discussed in detail elsewhere).38 Indeed, survival of patients with RA who had poor functional status over 5 years (Fig. 1A) was comparable to stage IV Hodgkin’s disease at that time at Stanford University (Fig. 1C),38 and three-vessel coronary artery disease at the Cleveland Clinic (Fig. 1D).39 This information suggested that patient questionnaire scores were not a poor surrogate for traditional measures, and, in fact, were as “scientific” as any data for the clinical prognosis of RA. Patient questionnaire scores, which had been used in the author’s clinic, then were transformed from a somewhat optional to a mandatory component of all visits in the author’s clinical care.

A prospective study, initiated in 1985, indicated that the MHAQ and multidimensional HAQ (MDHAQ), derived from the HAQ, predicted 5-year mortality at higher levels of significance than laboratory tests, radiographs, or joint counts in the same patients.40 Analysis of survival according to whether rheumatoid factor was present or absent indicates essentially no difference between the two groups (Fig. 2). By contrast, significant differences in survival were seen according to physical function; all patients with a normal physical function score of 0 (on a scale of 0-3) survived the 5-year period, while survival was only 65% among patients with the poorest physical function scores (Fig. 2).

All clinical measures at baseline, in 1985, except pain scores, indicated poorer status in patients who died prior to 1990, compared to survivors in 1990 (Table 2).40 However, higher levels of significance were seen for patient questionnaire scores and measures of functional status, compared to laboratory tests and radiographic scores (Table 2).40 Cox proportional hazard models indicated that the three significant independent variables which were prognostic of mortality were age, comorbidities, and functional status according to an MHAQ (Table 3). Thus, the functional status questionnaire score was significant, while joint counts, laboratory tests, and radiographic scores were not significant in multivariable analysis of clinical predictors of mortality.

Similar results have been reported in 16 additional studies (total of 18) in a descriptive analysis of all reports concerning prediction of mortality in patients with RA over periods of 5 years or longer (Fig. 3).41 Physical function was significant in 17 of 18 studies, 14 (72%) in multivariate analyses and three...
5-Year Survival in 206 Patients with RA: Cohort #2 – 1985-1990

(22%) in univariate analyses only, compared to comorbidities in 65% and 30% of 23 studies, rheumatoid factor in 45% and 21% of 29 studies, extra-articular disease in 44% and 39% of 18 studies, ESR in 37% and 32% of 19 studies, socioeconomic status in 31% and 46% of 13 studies, formal joint count in 22% and 28% of 18 studies, and a hand radiograph in 11% and 50% of 18 studies. An elevated score for physical function might be viewed as a marker of a dysregulation, comparable to elevated blood pressure, glucose, or cholesterol. Control of dysregulation prevents damage to organs such as blood vessels, kidneys, and joints, leading to improved outcomes including survival.

Over the last 25 years, extensive evidence has emerged which indicates that patient self-report questionnaires provide quantitative data of comparable “scientific” value to laboratory tests or joint counts in clinical care, as described in several previous essays and as summarized briefly below:

1. Patient questionnaire scores distinguish active from control treatment in clinical trials as efficiently as formal joint counts and laboratory tests

The relative efficiencies of the three patient-reported outcome measures in the Core Data Set – physical function,

Table 2 Mean Baseline Values in 206 RA Patients According to Survival or Mortality Status 5 Years Later*

<table>
<thead>
<tr>
<th>Measures</th>
<th>Alive</th>
<th>Dead</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>55.1</td>
<td>65.5</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>ARA functional class</td>
<td>2.2</td>
<td>2.6</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Number of comorbidities</td>
<td>1.1</td>
<td>2.1</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Walking time</td>
<td>10.8</td>
<td>16.8</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>ESR</td>
<td>33.8</td>
<td>48.3</td>
<td>0.004</td>
</tr>
<tr>
<td>MHAQ physical function score</td>
<td>1.98</td>
<td>2.32</td>
<td>0.005</td>
</tr>
<tr>
<td>Learned helplessness</td>
<td>2.41</td>
<td>2.55</td>
<td>0.007</td>
</tr>
<tr>
<td>Patient global estimate of status</td>
<td>2.6</td>
<td>3.0</td>
<td>0.01</td>
</tr>
<tr>
<td>Number of extra-articular features</td>
<td>0.2</td>
<td>0.5</td>
<td>0.02</td>
</tr>
<tr>
<td>Duration of disease</td>
<td>9.1</td>
<td>12.7</td>
<td>0.03</td>
</tr>
<tr>
<td>Years of education</td>
<td>10.8</td>
<td>9.4</td>
<td>0.03</td>
</tr>
<tr>
<td>Joint count</td>
<td>12.8</td>
<td>15.9</td>
<td>0.04</td>
</tr>
<tr>
<td>Radiographic score</td>
<td>1.2</td>
<td>1.4</td>
<td>0.20</td>
</tr>
<tr>
<td>Rheumatoid factor titer</td>
<td>2.7</td>
<td>2.9</td>
<td>0.28</td>
</tr>
<tr>
<td>Pain</td>
<td>5.40</td>
<td>5.19</td>
<td>0.68</td>
</tr>
</tbody>
</table>

*Mean baseline values, in 1985, in 206 patients with RA, according to whether patients were alive or dead 5 years later in 1990. Note that all measures indicated poorer status in patients who would not survive the 5-year period, except pain scores, but higher levels of significance were seen for patient questionnaire scores and measures of functional status, compared to laboratory tests and radiographic data (From Callahan, et al. Arthritis Care Res. 1997;10:381-94.) ARA, American Rheumatism Association; ESR, erythrocyte sedimentation rate; MHAQ, modified health assessment questionnaire.
2. Patient questionnaire scores are more reproducible than formal swollen and tender joint counts, and can be used by any physician

A careful joint examination clearly is required for diagnosis, and a formal tender and swollen joint count is the most specific indicator of change in RA status. However, joint counts are poorly reproducible in formal studies and are less sensitive in detecting inflammatory activity than magnetic resonance imaging (MRI) or ultrasound. Even when optically reliable, differences between observers have led to a requirement that the joint count be performed by the same observer from one visit to the next. Therefore, quantitative data concerning patient status cannot be collected reliably by other observers or general physicians, a serious limitation of the joint count for “scientific” patient care.

3. Patient questionnaire scores are more likely to reflect an abnormal state than laboratory tests

Patient questionnaire scores have been found abnormal in 80% to 90% of patients with RA or in all patients seen in a rheumatic disease clinic. By contrast, rheumatoid factor is positive in 71% and anti-CCP in 69% of patients with RA, while approximately 60% have abnormal ESR or CRP. Therefore, patient questionnaire scores are more frequently abnormal than laboratory tests, although, as noted above, a tender or swollen joint count is the most sensitive measure of disease activity.

Table 3  Cox Proportional Hazards Model Analyses of Survival Between 1985 and 1990

<table>
<thead>
<tr>
<th>Variable</th>
<th>RR (95% CL)</th>
<th>P Value</th>
<th>RR (95% CL)</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>1.07 (1.04–1.11)</td>
<td>&lt; 0.001</td>
<td>1.06 (1.03–1.10)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Comorbidity</td>
<td>1.63 (1.32–2.00)</td>
<td>&lt; 0.001</td>
<td>1.40 (1.11–1.77)</td>
<td>0.02</td>
</tr>
<tr>
<td>MHAQ physical function score</td>
<td>2.00 (1.28–3.12)</td>
<td>0.003</td>
<td>1.76 (1.40–2.78)</td>
<td>0.02</td>
</tr>
<tr>
<td>Disease duration</td>
<td>1.04 (1.01–1.06)</td>
<td>0.02</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Education</td>
<td>0.89 (0.82–0.97)</td>
<td>0.007</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>ESR</td>
<td>1.01 (1.00–1.02)</td>
<td>0.005</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Joint count</td>
<td>1.02 (0.97–1.04)</td>
<td>0.10</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Walking time</td>
<td>1.03 (1.01–1.06)</td>
<td>0.04</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Radiograph</td>
<td>1.04 (0.86–2.27)</td>
<td>0.17</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

*Baseline demographic, joint count, radiographic, laboratory, functional, questionnaire, and disease variables were entered into a model to analyze survival in 206 patients over the subsequent 5 years (From Callahan, et al. Arthritis Care Res. 1997;10:381-94.)

† 95% CL = 95% confidence limit. MHAQ, modified health assessment questionnaire; ESR, erythrocyte sedimentation rate.
sensitive measure in RA.

4. Treatment with placebo is more likely to result in improvement in joint counts, compared Patient questionnaire scores

Among the seven ACR Core Data Set measures, patients who receive placebo or control treatment in clinical trials generally show improvement in swollen and tender joint counts, but are much less likely to show improvement in patient questionnaire measures and laboratory tests.51,52,68,69

5. Patient questionnaire scores are correlated significantly with traditional joint counts, radiographs, and laboratory tests

Significant correlations of patient questionnaire scores with joint counts, laboratory tests, and radiographic scores were documented in the 1980s,43 albeit with different levels and with recognition of two categories of measures (Fig. 4). The first category includes traditional “objective” measures: radiographs, joint deformities, and laboratory tests, including histocompatibility type, which are correlated at higher levels with one another than with patient questionnaire scores or tender joint counts.70 The second category includes patient questionnaire scores and tender joint counts, which are correlated at higher levels with one another than with radiographic or laboratory measures.43 Laboratory and radiographic measures clearly are more important in pathogenesis and more specific to RA. However, patient self-report measures are far more significant in clinical prognosis than radiographs or laboratory measures, as noted below, and, therefore, are of comparable “scientific” value in clinical care.

6. Patient questionnaire scores for physical function are far more prognostic for severe long-term RA outcomes, including work disability and mortality, than radiographic or laboratory data

All studies that include a patient questionnaire indicate that a baseline questionnaire measure of physical function is far more significant than a radiograph or laboratory test in the prognosis of severe long-term outcomes of RA, including work disability, costs, joint replacement surgery, and death (all except radiographic damage).71

7. An MDHAQ/RAPID3 score is informative in patients with all rheumatic diseases

Most patients with any rheumatic disease experience problems in physical function, pain, or global status, quantified by MDHAQ/RAPID3 scores, as well as morning stiffness and fatigue, as assessed on the MDHAQ.72 MDHAQ is informative in patients with all rheumatic diseases.73

8. RAPID3 can be calculated in fewer than 10 seconds, compared to 90 seconds for a swollen and tender joint count and 2 minutes for a CDAI or DAS28

As noted, a RAPID3 score can be calculated in fewer than 10 seconds.74 By contrast, performance of a swollen and tender 28-joint count requires about 90 seconds. Calculation of a DAS28 or CDAI requires at least 2 minutes (including the

---

**Figure 4** Two types of measures to assess RA. All measures are correlated significantly, including those from radiographs, joint examination, laboratory tests, and patient questionnaire scores.70 However, two clusters of quantitative measures appear.43 The first cluster (left circle) includes traditional “objective” measures: radiographs, joint deformities, and laboratory tests, and histocompatibility type, which are correlated at higher levels with one another than with patient questionnaire scores or tender joint counts.70 The second cluster (right circle) includes patient questionnaire scores and tender joint counts, which are correlated at higher levels with one another than with radiographic or laboratory measures.43 Laboratory and radiographic measures clearly are more important in pathogenesis and more specific to RA. However, patient self-report questionnaire measures are far more significant in prognosis of severe outcomes, including work disability and mortality, than radiographs or laboratory measures, and therefore of comparable “scientific” value in clinical care.
joint counts), even when all the data are readily available. The advantage of specificity is limited by limitations of feasibility in usual care.

9. Treatment guided by quantitative data results in better patient status than usual nonquantitative clinical care

Six clinical trials have now documented that guidance using quantitative data results in better patient status than usual care without such guidance: the Finnish Rheumatoid Arthritis Combination Therapy (FIN-RACo) trial; Tight Control for Rheumatoid Arthritis (TICORA) trial; Be- handel Strategien (BeSt) or “treatment strategies” trial; Computer Assisted Management in Early Rheumatoid Arthritis (CAMERA) study; the Ciclosporine, Methotrexate, Steroid in Rheumatoid Arthritis (CIMESTRA) trial; and the TICORA2 trial. All six trials used the DAS28 to generate quantitative data. DAS28 is not available at many visits in usual clinical care, and RAPID3 is correlated significantly with DAS28. Therefore, a RAPID3 score could be appropriate for usual clinical care, although prospective studies are needed to confirm this hypothesis.

Conclusions

Laboratory tests are the most technologically advanced measures, and joint counts are the most specific measures for RA. However, patient self-report questionnaire scores detect treatment effects at similar levels, are more significant in prognosis, and considerably more feasible and cost-effective than laboratory tests or joint counts. A patient questionnaire may be regarded as a quantitative patient history, equally “scientific” to monitor patient status as laboratory tests and joint counts.

A patient questionnaire does not replace a further patient history, joint counts, or laboratory tests. RAPID3 scores, based on self-report patient questionnaire scores, provide informative quantitative data for patient status from one visit to the next. If quantitative data are recorded, an opportunity for documentation and more rational monitoring is gained, along with enhanced efficiency of patient care. If no data are recorded, this opportunity is lost and can never be replaced.

It has been suggested that “80% of the data in 100% of the patients may be preferable to 100% of the data in 5% of the patients” (or fewer) who might be included in clinical research. Therefore, a less comprehensive measure, which is feasible and applicable in usual clinical care, appears preferable to no quantitative measure at all. However, a RAPID3 score may provide more than “80%,” and indeed may be as informative as a DAS or CDAI for patient assessment, reflecting patient and physician goals of treatment as accurately as the number of swollen and tender joints. All rheumatologists would find it valuable to ask all patients at all visits to complete a MDHAQ and for a staff member or themselves to score a RAPID3 in 5 seconds in usual care.

Disclosure Statement

Dr. Pincus has received consultant fees and lecture fees from Bristol-Myers Squibb. He has received consultant fees from and served on a Scientific Advisory Board for UCB, Inc. He has received research support from Centocor, Inc.

References

18. Sokka T, Pincus T. Ascendancy of weekly low-dose methotrex-
27. Sokka T, Pincus T. Erythrocyte sedimentation rate, C-reactive protein, or rheumatoid factor are normal at presentation in 35%-45% of patients with rheumatoid arthritis seen between 1980 and 2004: analyses of patients from Finland and the United States. J Rheumatol. 2009;36(7):1387-90.
46. Pincus T, Wolfe F. Patient questionnaires for clinical research and improved standard patient care: is it better to have 80% of the information in 100% of patients or 100% of the information in 5% of patients? J Rheumatol. 2005;32:575-7.
47. Pincus T. The DAS is the most specific measure, but a patient questionnaire is the most informative measure to assess rheumatoid arthritis. J Rheumatol. 2006;33:834-7.
51. Strand V, Tugwell P, Bombardier C, et al. Function and


67. Pincus T, Swearingen CJ. Erythrocyte sedimentation rate (ESR) is the least likely of Core Data Set measures to identify an “abnormal state” in new patients with RA to monitor therapeutic responses, according to 3 definitions of “abnormal state.” (Abstract #318). Arthritis Rheum. 2009;60(Suppl):S117.


73. Pincus T, Sokka T. Can a Multi-Dimensional Health Assessment Questionnaire (MDHAQ) and Routine Assessment of Patient Index Data (RAPID) scores be informative in patients with all rheumatic diseases? Best Pract Res Clin Rheumatol. 2007;21(4):733-53.


78. Gokoop-Ruiterman YPM, de Vries-Bouwstra JK, Allaart CF,


