NEW CLASSIFICATION CRITERIA FOR RHEUMATOID ARTHRITIS AND ANKYLOSING SPONDYLITIS: GAME CHANGERS OR MERE NAME CHANGERS?

Rohini Handa, New Delhi

ABSTRACT
The current treatment paradigm for inflammatory rheumatic diseases like rheumatoid arthritis (RA) and ankylosing spondylitis (AS) revolves around early diagnosis and tight disease control. There is ample evidence to show that treatment instituted earlier in the disease course leads to better outcomes. Consequently, the older classification criteria, derived from patients with typical clinical features of well-established disease, are undergoing revision. The emphasis is on early recognition without sacrificing sensitivity for specificity. This chapter reviews the recently proposed classification criteria for RA and AS and puts in perspective the significance of these changes for an internist.

INTRODUCTION
Rheumatology as a specialty has an abundance of criteria of several different types: classification criteria, prognostic criteria, diagnostic criteria, remission criteria, etc. Criteria, while being an unavoidable necessity in this era of evidence-based medicine, can be quite daunting for the uninitiated. To add to the complexity, criteria do change as knowledge advances. The need to understand that classification criteria evolved in response to the fact that most systemic rheumatic diseases lack a single pathognomonic or distinguishing feature. Features like arthritis, morning stiffness, fever, Raynaud’s phenomenon, skin rash are common to several diseases. Also, many of the features do not occur concurrently but sequentially. These features present in a particular combination along with certain laboratory investigations help identify a specific disease. Classification criteria help differentiate patients with a particular disease from patients with a potentially confusing condition as well as from the normal population (case definition). For example, criteria for AS would help differentiate AS from other causes of back pain. The issue that classification criteria tend most often to answer is not whether the patient has a disease, but which disease does the patient have. Classification criteria serve several useful purposes: provide uniformity for patients being included in epidemiological studies and trials, facilitate comparison between different centers and provide a lingua franca for scientific communication.

The clinician would do well to remember that classification criteria, despite being (mis)used as surrogate diagnostic criteria, are not meant (and were never meant) to be used for diagnostic purposes. This difference is crucial. ‘Classification’ criteria are applicable to groups and are more specific than ‘diagnostic’ criteria which are applied to individuals and are more sensitive. Diagnosis is classification in the individual patient. While using criteria, the clinician should never lose sight of the fact that therapeutic decisions in an individual patient should not be governed solely by fulfillment or lack of fulfillment of criteria.

In this write-up I shall confine myself to the new classification criteria for RA and AS and their clinical implication.

RHEUMATOID ARTHRITIS
Historical perspective
RA is the commonest inflammatory polyarthritis seen in clinical practice. The first attempt to
classify RA by the American Rheumatism Association (now, the American College of Rheumatology) was way back in 1956. These criteria incorporated 11 clinical, serological, radiological and histological features with 19 exclusions. These sets classified disease as ‘definite’, ‘probable’ and ‘possible’ RA. The ‘definite’ group required 5 criteria with 6 weeks of symptoms while the ‘probable’ RA required 3 criteria with 4 weeks of joint symptoms. The premise was that in the ‘definite’ group there should be almost no question that every included patient has RA while in the ‘probable’ group the likelihood should be great that every included patient has RA. The criteria for the ‘possible’ group were much less rigid and it was conceded that some patients who do not have RA would get included in this category. The ethos was to pick up early and atypical cases which can be followed profitably and at no time can this group be used for definite conclusions concerning the characteristics or treatment of RA. Even the first criteria underscored the important fact that these criteria were not designed to aid a physician in making a diagnosis in an individual case, but rather to establish the findings necessary to allow inclusion of a patient in one or another of the categories. The 1958 revision added the category of ‘classic’ RA for patients exhibiting 7 out of 11 criteria. The authors did mention that errors in diagnosis will become apparent as the patients in possible category are followed, and at no time can this group be used for definite conclusions concerning the characteristics or treatment of RA. Even the first criteria underscored the important fact that these criteria were not designed to aid a physician in making a diagnosis in an individual case, but rather to establish the findings necessary to allow inclusion of a patient in one or another of the categories. The 1958 revision added the category of ‘classic’ RA for patients exhibiting 7 out of 11 criteria.4

One lacuna of these criteria was the inability to pick up disease during remission. The other drawback was inclusion of synovial fluid analysis and biopsy (synovial or rheumatoid nodule), invasive procedures never done in many patients. These lacunae were corrected in the Rome criteria where poor mucin clot and histological features of synovial and nodule biopsy were dropped from the criteria. The Rome criteria also differentiated active from inactive disease.5 Another adaptation, the New York criteria never gained much acceptance because it was cumbersome.6

1987 Criteria for RA

The criteria for RA were revised after a long gap of 25 years in 1987.7 The 1987 ARA criteria were developed using RA cases and controls attending hospital clinics. The patients included had longstanding disease (mean disease duration 7.7 years). These criteria (Table 1) incorporated the typical features of symmetric inflammatory polyarthritis and did away with the categories of definite, possible and probable. These criteria were simple to use and required only one laboratory test, rheumatoid factor, and only one set of radiographs, posteroanterior view of hands and wrists. The 1987 criteria were widely adopted all over the world and paved the way for uniformity in case inclusion. These criteria had a sensitivity of 91-94% and specificity of 89% when comparing RA with non-RA. These criteria served their purpose admirably well for several years.

Over a period of time a few shortcomings became apparent. The first was the poor performance characteristics of 1987 criteria in early RA. This coincided with a shift in the focus in RA from ‘established’ to ‘early disease’. Two things have fuelled interest in early RA: an explosion of targeted biologic therapies and the growing realization that time to treat is a key driver of outcome. The cut offs between ‘early’ and ‘established’ RA have progressively decreased. Currently, ‘early RA’ is defined as disease duration less than 1 year. Some authorities split this into ‘very early RA’ with disease duration less than 3 months and ‘late early RA’ when the duration of symptoms ranges from 3-12 months. When applied to early inflammatory polyarthritis, the 1987 ARA criteria for RA had a low ability to discriminate between patients who developed persistent, disabling, or erosive disease and those who did not. Studies have shown that the 1987 ACR criteria, when applied to early RA, have a sensitivity ranging from 40 to 90% and specificity from 50 to 90%.8-12

The second drawback of 1987 criteria was the inclusion of radiographic features. The radiologic criterion of erosions is encountered in a very small proportion (~13%) of patients in the first 3 months of disease onset limiting its utility.13 However, as many as 50-70% patients may have erosive disease by 2 years thereby underscoring the importance of early treatment.14

Over the past decade anti-citrullinated peptide antibodies (ACPA), also known as anti-cyclic citrullinated peptide antibodies (anti-CCP), have emerged as an important serologic marker for RA. These predict erosive disease and are poor prognostic markers.15 These antibodies obviously do not find mention in the 1987 criteria which were formulated prior to the advent of ACPA. Liao and colleagues attempted to incorporate ACPA in the 1987 criteria which improved upon the sensitivity of the ACR criteria, most remarkably for subjects with symptoms <6 months.16 However, their attempt did not attract widespread attention. The latest attempt in remission, the 2010 criteria, aim to rectify many of these shortcomings.

2010 ACR/EULAR Criteria for RA

The 2010 criteria emerged as a joint initiative of American and European workers and were published simultaneously.
The performance characteristics of the 2010 ACR/EULAR criteria are being tested in early arthritis. Alves et al. applying these criteria to patients with early arthritis (<12 months) in the Rotterdam Early Arthritis Cohort (REACH) reported a sensitivity of 70% and equal specificity (70%). Using the cut-point of 6 to start treatment, in this study 30% of persistent patients would not be treated, whereas 30% of the non-persistent patients would have been. Lowering the cut-point to 4 increased sensitivity to 0.92 at the cost of specificity (0.33). Increasing it to 7 had a sensitivity of 0.53 and a specificity of 0.85. Similar results have been reported by other groups. Of the 301 patients with early arthritis (0-12 months), only 28% fulfilled the 1987 ACR criteria at baseline while 45% satisfied the 2010 ACR/EULAR criteria for RA. Nearly one-third of the patients in this cohort who would have been labeled as undifferentiated arthritis using the 1987 criteria met the 2010 criteria for RA.

Overdiagnosis, however, remains an area of concern with the new criteria. More patients whose disease eventually resolved without ever requiring DMARD were classified at baseline as RA according to the 2010 criteria than with the 1987 criteria (8%) vs 2%; p=0.01) in a study from UK. Similar concerns have been voiced from India especially in context of infectious arthritis where patients with Chikungunya arthritis may easily satisfy the 2010 ACR/EULAR criteria for RA. Clearly, the quest for ideal criteria is far from over. The 2010 criteria for RA represent a significant advance but may need refinement in different populations and as new knowledge becomes available.

ANKYLOSING SPONDYLITIS (AS)

General considerations

AS is the prototypic illness of spondarthritides, a group of interrelated disorders characterized by inflammatory low back pain and the presence of HLA B-27. The commonality of low back pain (LBP), said to be the second most frequent ailment of mankind after common cold, makes the task of the clinician even more difficult. Diagnostic delays of 5-10 years are not uncommon. These delays are more pronounced in women. A diagnostic delay of 6.9 ±5.2 years has been reported in a series of 70 patients from India. The main

### Table 2: The 2010 ACR/EULAR classification criteria for rheumatoid arthritis

<table>
<thead>
<tr>
<th>Domain: Joint involvement</th>
<th>Domain: Duration of synovitis</th>
</tr>
</thead>
<tbody>
<tr>
<td>• 1 large joint (0 points)</td>
<td>• Less than 6 weeks (0 points)</td>
</tr>
<tr>
<td>• 2-10 large joints (1 point)</td>
<td>• 6 weeks or longer (1 point)</td>
</tr>
<tr>
<td>• 1-3 small joints (2 points)</td>
<td>• Abnormal ESR/CRP (1 point)</td>
</tr>
<tr>
<td>• 4-10 small joints (3 points)</td>
<td>• Normal ESR/CRP (0 points)</td>
</tr>
<tr>
<td>• &gt;10 joints [at least 1 small joint] (5 points)</td>
<td>• RF or CCP negative (0 points)</td>
</tr>
<tr>
<td></td>
<td>• RF or CCP positive at low titer, &lt;3 times ULN (2 points)</td>
</tr>
<tr>
<td></td>
<td>• RF or CCP positive at high titer, defined as &gt;3 times ULN (3 points)</td>
</tr>
</tbody>
</table>

Domain: Serology

<table>
<thead>
<tr>
<th>Domain: Acute phase reactants</th>
</tr>
</thead>
<tbody>
<tr>
<td>• RF or CCP negative (0 points)</td>
</tr>
<tr>
<td>• RF or CCP positive at low titer, $&lt;3$ times ULN (2 points)</td>
</tr>
<tr>
<td>• RF or CCP positive at high titer, defined as $&gt;3$ times ULN (3 points)</td>
</tr>
</tbody>
</table>

ULN= upper limit of normal

Joint involvement refers to any swollen or tender joint on examination, which may be confirmed by imaging evidence of synovitis. Distal interphalangeal joints, first carpometacarpal joints, and first metatarsophalangeal joints are excluded from assessment.

“Large joints” refers to shoulders, elbows, hips, knees, and ankles.

“Small joints” refers to the metacarpophalangeal joints, proximal interphalangeal joints, second through fifth metatarsophalangeal joints, thumb interphalangeal joints, and wrists.

in the ACR and EULAR journals. The major aim was to permit early identification of poor prognosis arthritis much before the classic features of florid disease became apparent. These criteria, listed in Table 2, are meant to be applied to patients newly presenting with undifferentiated inflammatory synovitis. These incorporate factors that best discriminate between those patients who are and those who were not at high risk for persistent and/or erosive disease—this being the appropriate current paradigm underlying the disease construct ‘RA’. The gold standard for diagnosis of RA was methotrexate initiation by the physician. This was used to identify clinical and laboratory variables which were then subjected to consensus-based, decision science informed approach leading to the evolution of a scoring system. The criteria were satisfied in 87-97% of the patients where physicians instituted methotrexate.

In the new criteria set, classification as ‘definite RA’ is based on the confirmed presence of synovitis in at least one joint, absence of an alternative diagnosis better explaining the synovitis, and achievement of a total score of 6 or greater (of a possible 10) from the individual scores in four domains: number and site of involved joints (range 0–5), serological abnormality (range 0–3), elevated acute-phase response (range 0–1) and symptom duration (two levels; range 0–1) [Table 2]. These criteria have done away with features that are typical of late disease, namely symmetry, rheumatoid nodules and radiographic changes. There is no longer insistence on disease duration of 6 weeks. The criterion of morning stiffness has been dispensed with and the serologic marker of ACPA included. This practically means that a patient with 1 small joint involvement (2 points), high levels of RF/ACPA (3 points) and high ESR/CRP (1 point) can be classified as RA even on day 1 of symptoms. These criteria have a provision whereby some patients can be classified as RA even if they do not fulfill the criteria. These include patients with erosive disease typical of RA with a history compatible with prior fulfillment of the 2010 criteria and patients with long-standing disease, including those whose disease is inactive (with or without treatment) who, based on retrospectively available data, have previously fulfilled the 2010 criteria.

The criteria for RA represent a significant advance but may need refinement in different populations and as new knowledge becomes available.
ASAS have proposed new ASAS-IBP criteria which discriminate IBP from non-IBP, a subtle but important difference.26 The ASAS criteria for inflammatory back pain to be applied to patients with chronic back pain >3 months include 5 parameters: age at onset <40 years, insidious onset, improvement with exercise, no improvement with rest and pain at night (with improvement upon getting up). The criteria, fulfilled if at least four out of five parameters are present, have a sensitivity of 79.6% and specificity of 72.4%. The ASAS-IBP criteria have been shown to have lesser sensitivity but better specificity than the Calin criteria (sensitivity 89.9% and specificity 52.5%) in the ASAS validation study.

The major bottleneck in early recognition of AS has been the insistence on radiologic evidence of sacroiliitis in most criteria: the Rome criteria proposed in 1961 and published in 1963 [Reference 27], the New York criteria for AS first proposed in 1966 and published in 1968 [Reference 6], and modified in 1984 and the Mau et al criteria for determining early AS proposed in 1990.6,27-29 The widely employed modified NY criteria for AS are listed in Table 3. These criteria focus on AS. The European Spondyloarthropathy Study Group (ESSG) criteria (Figure 1) expanded the concept to include other subsets of spondarthritides.30

The modified New York criteria are unable to detect patients with early AS because of the radiologic criterion of sacroiliitis. In the olden days this may not have mattered because treatment options for AS were limited. With the advent of TNF inhibitors and the opening of new treatment vistas early diagnosis is imperative. HLA B-27 does not find mention in these criteria. Another drawback of the NY criteria is that they focus on axial disease only. The spectrum of SpA includes peripheral disease, axial disease and both axial plus peripheral disease. The Amor and European Spondyloarthropathy Study Group (ESSG) criteria include indeed the entire SpA spectrum, but do not specify the predominant feature (axial or peripheral).

**ASAS Criteria for SpA**

The ASAS have proposed classification criteria for axial disease in 2009 and peripheral SpA in 2011.31,32 These permit identification of both early and established cases and...
ASAS classification criteria for axial spondyloarthritis (SpA) in patients with ≥ 3 months back pain and age at onset <45 years

<table>
<thead>
<tr>
<th>Sacroiliitis on imaging* plus ≥1 SpA feature*</th>
<th>Arthritis or Enthesitis or Dactylitis</th>
<th>or</th>
</tr>
</thead>
<tbody>
<tr>
<td>HLA-B27 plus ≥2 other SpA features*</td>
<td>plus ≥ 1 of:</td>
<td>or</td>
</tr>
<tr>
<td>• elevated CRP</td>
<td>- Psoriasis</td>
<td>plus ≥ 2 of the remaining:</td>
</tr>
<tr>
<td>• HLA-B27</td>
<td>- Inflammatory bowel disease</td>
<td>- Arthritis</td>
</tr>
<tr>
<td>≥ 1 SpA feature</td>
<td>- Preceding infection</td>
<td>- Enthesitis</td>
</tr>
<tr>
<td>• HLA-B27</td>
<td>- HLA-B27</td>
<td>- Dactylitis</td>
</tr>
<tr>
<td>• elevated CRP</td>
<td>• definite radiographic sacroiliitis according to mod NY criteria</td>
<td>- IBP in the past</td>
</tr>
<tr>
<td>• positive family history for SpA</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*SpA features
- inflammatory back pain
- arthritis
- enthesitis (heel)
- uveitis
- dactylitis
- psoriasis
- Crohn’s/colitis
- good response to NSAIDs
- family history for SpA
- HLA-B27
- elevated CRP
- ≥ 1 SpA feature
- ≥ 2 other SpA features

Fig. 2: ASAS Classification Criteria for axial SpA

incorporate MRI detection of sacroiliitis (Figures 2 and 3). HLA B27 also finds mention in the new criteria. Another significant important difference from the older criteria is that inflammatory back pain (IBP) is not considered obligatory. The entry criterion is back pain of at least 3-months with onset before 45 years. IBP has been included as an extra SpA feature.

Plain x-rays and even CT scans can detect structural changes like joint margin irregularity and erosions but not inflammation. The major advance that has permitted early recognition of AS is the availability of MRI. T2-weighted sequences suppress the signal from bone marrow fat allowing visualization of the bright signal from free water reflecting inflammation in subchondral bone marrow of the sacroiliac joints, vertebrae and posterolateral elements of the spine. The recommended MRI protocol is T1-weighted turbo spin-echo sequence, a T2-weighted gradient-echo sequence using the opposed-phase technique and a STIR sequence with slices of 4 mm thickness. For diagnosis, MRI of SI joints is obtained while MRI of spine is more frequently used to assess the activity of the disease and to assess efficacy of treatment with TNF inhibitors.

The new ASAS criteria incorporate the predominant feature—axial or peripheral. The sensitivity and specificity of axial SpA criteria are 82.9% and 84.4%, respectively. The criteria for peripheral SpA are applicable to patients with peripheral arthritis (usually predominantly of the lower limbs and/or asymmetric arthritis), and/or enthesitis, and/or dactylitis. The criteria have a sensitivity of 77.8% and specificity of 82.2%. The criteria for axial SpA and for peripheral SpA can be combined to get one criteria set for SpA: patients with back pain (with or without peripheral symptoms) follow the criteria set for axial SpA; patients with peripheral symptoms only follow the criteria for the peripheral SpA. This combined set has a sensitivity and specificity of 78.0% and 83.7%, respectively.

CONCLUSIONS

From the foregoing discussion it is clear that the science of classification is a science in evolution. Criteria will continue to be refined with the availability of new knowledge. The challenge is to blend current classification criteria into existing patient care algorithms without sacrificing sensitivity or specificity or adding to the complexity— a difficult but not unachievable goal!

REFERENCES


