Lessons from an International Survey of Paper Cases of 10 Real Patients from an Early Arthritis Clinic

JEAN-MARIE BERTHELOT, METTE KLARLUND, DENNIS MCGONAGLE, HEIN J. BERNELOT-MOENS, ANDREI CALIN, BEVERLEY HARRISON, H. RALPH SCHUMACHER, KALEVI KAARELA, ALEXANDROS A. DROSOS, JAN-L. HÜLESEMMANN, WEI-HOWE KOH, YRJO-T. KONTTINEN, LEONARDO PUNZI, KIYOAKI TANIMOTO, H. JAMES WILLIAMS, FREDERICK WOLFE, CRISTIANO-A. ZERBINI, ALAIN SARAUX and the CRI (Club Rhumatismes et Inflammation) Group

ABSTRACT.

Objective. To determine how experts would classify 10 early-arthritis cases (7 atypical) and to study discrepancies in diagnoses relative to ACR criteria for rheumatoid arthritis (RA) or ESSG criteria for spondyloarthropathy (SpA).

Methods. Ten real cases (5 met ACR criteria for RA, 6 ESSG criteria for SpA, 3 both and 2 neither) followed for 28.5 ± 4.8 months were sent as paper cases to 20 international and 12 French experts. Each expert selected a diagnosis among 8 possible choices and rated it on a 0-10 confidence scale. For each case, 3 analog scales (0-100 mm) were used to indicate the probability of RA, SpA or undifferentiated arthritis (UA).

Results. Experts often disagreed about diagnoses (up to 5 different diagnoses for a given case, with a mean of 3.9 per case). Similarly, expert opinions on probabilities for RA and SpA differed widely, with great overlap between confidence for RA, SpA and UA. Fulfilment of ACR or ESSG criteria was poorly related to the experts' diagnosis and evaluation of probabilities for RA and SpA. However, UA was a relatively infrequent choice (19%).

Conclusions. There was no general consensus about the nosology of early RA and SpA. Classification of atypical early arthritis was not resolved by currently available criteria for RA and SpA. This may have implications for therapy in early disease. (J Rheumatol 2001;28:975-81)

Key Indexing Terms:

EARLY ARTHRITIS
RHEUMATOID ARTHRITIS
CRITERIA
SPONDYLOARTHROPATHY
CONFIDENCE
DIAGNOSIS

Although early aggressive treatment of rheumatoid arthritis (RA) is now advocated by many experts, this approach can be expensive and/or produce occasional but serious side effects. Hence, a high level of confidence in the diagnosis of RA early in its course may be mandatory. Classification criteria have been constructed for epidemiological or research studies to differentiate patients with well-established RA from those with other forms of long
lasting rheumatic conditions\textsuperscript{7-8}. However, these criteria are sometimes improperly used as diagnostic tools for early arthritis. In fact, there is no guarantee of their validity in this context since clinical features of early RA and other forms of rheumatism could overlap more often than in well-established disease. The distinction between early RA and undifferentiated arthritis (UA) or early spondyloarthropathy (SpA) is much finer than that between destructive RA and ankylosing spondylitis. Moreover, atypical early arthritis is not a rare event: in recent studies on the outcome of early arthritis, 22 (\(73/332\))\textsuperscript{9}, 32 (\(75/233\))\textsuperscript{10}, and even 53\% (\(512/966\))\textsuperscript{11} of patients remained unclassified after one year of observation. Similarly, in a 3 year follow-up study of 270 patients with early arthritis begun in 1995 in the West of France (Brittany), the diagnosis remained unclear in \(61/270\) (23\%), and changed between the first and last examination in 96 of the remaining 209 cases (46\%). RA and European Spondylarthropathy Study Group (ESSG) criteria were recorded for all 270 patients. Seven of these 157 cases (\(61 + 96\)) were selected and sent as paper cases to 26 international experts who have published on early arthritis, together with 3 other cases typical of early RA, early SpA and early psoriatic arthropathy.

The first aim of our study was to determine how often experts choose to leave cases unclassified, the second to assess the overall confidence of experts in their diagnoses, the third to determine the extent of discrepancy among rheumatologists concerning atypical early arthritis, and the fourth to ascertain whether the experts' judgments about recent onset, atypical arthritis were strongly correlated with the fulfilment of American College of Rheumatology (ACR) criteria for RA or ESSG criteria for SpA.

MATERIALS AND METHODS

Selection and description of the 10 cases. Ten real cases from a cohort of 270 patients from 7 early arthritis clinics in France, followed up for more than 2 years, were selected for this survey. Seven out of the 10 cases were chosen because office based rheumatologists had difficulty in classifying them (mainly due to hesitation between RA and SpA or between SpA and UA). Three cases from this cohort considered highly suggestive of early RA, early SpA, or early psoriatic arthropathy were added.

The 10 cases are presented in the Appendix 1. The descriptions of the 10 cases were sufficiently explicit, so that each expert could easily check whether the patients satisfied the 1987 criteria for RA (format list, cumulatively) and ESSG criteria for SpA. In fact, 5 cases satisfied at least 4 ACR criteria for RA (1, 4, 5, 6, and 9) and 6 satisfied ESSG criteria for SpA (1, 2, 4, 5, 7, and 8). In 3 cases (1, 4 and, 5), both sets of criteria were fulfilled, although this was not indicated in the reports for these cases. Experts were not explicitly told to use either the tree or list format.

Each expert was asked to choose a diagnosis among a list of 8 for each of the 10 cases and indicate his degree of confidence on a scale of 0 to 10 (experts did not agree on terminology and description in advance). These 8 possible diagnoses were RA, SpA, psoriatic arthropathy (axial type), psoriatic arthropathy (peripheral type), ankylosing spondylitis (AS), reactive arthritis, UA, and "another diagnosis" (as specified). Among the 20 experts, a few diagnosed some patients as having both peripheral and axial arthropathies. One expert (No. 15) did not rate his confidence in the diagnoses.
The experts were also asked to estimate on a 0-100 analog scale the probability that these cases of early arthritis could be early RA, early SpA or UA. One expert (No. 20) did not reply to this second part of the survey.

Selection of the 20 experts. Twenty-six experts from 12 countries were consulted on the basis of their publications on early arthritis and/or their international reputation and/or their geographical location. All experts were from university centers. Twenty of the 26 experts (most between 40 and 60 years of age) agreed to reply. Twelve French experts also replied to the same questionnaire to determine whether the nosology of RA or SpA might differ in their country. All experts understood that their individual diagnoses would not be recognizable in the final report, and none were aware of the replies of the others. The 20 international experts who agreed to reply were from the U.S.A. (5), United Kingdom (3), Finland (2), Germany (2), the Netherlands (1), Denmark (1), Greece (1), Italy (1), Brazil (1), Singapore (1) and Japan (1). One remained unidentified as to country.

RESULTS

Discrepancies in diagnoses. At least 4 diagnoses were selected for each of the first 8 cases, partly because some experts chose 2 diagnoses (e.g. peripheral and axial psoriatic arthropathies). When patients diagnosed as SpA, reactive arthritis, psoriatic arthropathy, AS, and enteropathic arthritis were grouped together within the broad concept of spondyloarthropathies (SpA), discrepancies were less striking. However in 6/10 cases (Cases 1 to 6), patients were classified as RA by some experts and either SpA or UA by others (Figure 1).

Confidence in diagnoses. Despite a lack of agreement, confidence in the diagnoses was generally high for the 19/20 international experts who rated this factor, as illustrated by Case 1 (Figure 2). However, some experts were less confident than others. A rather high level of confidence was observed even in Cases 1 to 6 for which up to 6 different diagnoses were chosen by the 20 experts (Table 1).
Figure 2. Diagnoses selected by the 20 international experts for Case 1, with their level of confidence indicated on a 0 to 10 scale. SP: spondyloarthropathy.

Table 1. Experts' confidence in diagnosis for each of the 10 cases.

Discrepancies in estimations of RA probability. The probability that these cases might be early RA differed markedly, e.g. from 0 to 91% for case 1 (Figure 3).
Discrepancies in estimations of SpA probability. Marked discrepancies were also apparent concerning the probability of early SpA, e.g. from 0 to 90% for Case 1 (Figure 3).

Discrepancies in estimations of UA probability. Some experts were more reluctant than others to leave cases unclassified. Hence, the probability that cases could be UA also differed widely, e.g. from 0 to 92% for Case 1 (Figure 3).

Overlapping probabilities for RA, SpA and UA. A possible hesitation between RA and SpA is indicated by the differences in mean confidence for these diagnoses. There was a considerable overlap in the experts' evaluation of the probability of diagnosis of RA, SpA and UA for the more atypical cases (Figure 3). For Cases 2 and 4, the overlap was between SpA and UA (Figure 3). In the remaining cases (7 to 10) the diagnosis of SpA or RA was clearly more probable than for the others.
Probability of RA estimated by the 20 international experts compared to the fulfilment of ACR criteria for RA. Out of the 5 cases (1, 4, 5, 6 and 9), which satisfied (cumulatively) the 1987 criteria for RA (format list), only the most typical case (9) was diagnosed as RA by all experts. Case 1 was diagnosed as RA by only 6/20 experts, Cases 5 and 6 by only 3/20, and Case 4 by only 2/20 (Figure 1). The probability of these cases being RA was below 50/100 except for Case 9 (Figure 3). Conversely, cases that did not fulfil the 1987 ACR criteria were classified as RA by some experts: Case 2 (4/20 experts) and Case 3 (1/20 experts).

Comparison of the fulfilment of ESSG criteria for SpA and the experts' classifications. Out of the 6 cases (1, 2, 4, 5, 7 and 8) which satisfied (cumulatively) ESSG criteria for SpA, 4/6 were diagnosed as RA (instead of SpA) by some experts: Case 1 (6/20 experts), Case 2 (4/20 experts), Case 4 (2/20 experts), and Case 5 (3/20 experts), as well as the 3 cases that satisfied both RA and ESSG criteria (1, 4, and 5) (Figure 3). The association between the fulfilment of ESSG criteria for SpA and the mean probability that these 6 cases might be SpA was imperfect. The probability of these cases being SpA was above 50/100 for Cases 8, 4 (classified as RA by one expert) and 2 (classified as RA by 4 experts), but below 50/100 for Cases 7, 1 (classified as RA by 6 experts) and 5 (classified as RA by 3 experts) (Figure 3). Hence, ESSG criteria were poorly specific when the experts' classification was used as a gold standard within this panel of atypical early arthritis.

Comparison with the diagnoses of the 12 French experts. The replies of the 12 French experts were as diverse as those of the international panel, and their level of confidence in the diagnosis was quite similar. However, slight differences with the international panel were observed. For instance, for Case 5, French experts considered RA as a twice more probable diagnosis than UA, whereas the international panel found an equal probability for those 2 diagnoses. Conversely, French experts rated SpA as the most probable diagnosis in Cases 1 and 6, for which the international panel found RA and UA respectively. This is in accordance with the second slight difference, namely that French experts were more confident than the international panel about the possibility of SpA for 9/10 cases and even more reluctant than the international experts to consider patients as UA.

DISCUSSION

Our survey suggests that most of the 32 experts who replied to a mailed questionnaire about 10 cases of early arthritis were reluctant to admit that they should be considered UA. This is illustrated by the fact that UA was the most frequent choice for only one of the paper cases (Case 6; Figures 1 and 3). In addition, this survey indicates that some experts were more confident than others about their diagnoses: mean confidence on a 0-10 analog scale ranged from 5.6 ± 2.0 (Expert 17) to 8.7 ± 3.2 (Expert 20).

Our survey therefore confirms that discrepancies among rheumatologists confronted with atypical early arthritis may be considerable. Although very good agreement was observed for typical RA (Case 9), striking differences were noted for most of the other cases. For instance, 4 to 5 different diagnoses were selected by the 20 experts for each of the first 8 cases, and disagreement was still present when all diagnoses but RA and UA were grouped together. Moreover, despite these discrepancies, most experts rated their confidence as rather high (7 or more; Table 1), although sometimes for quite different diagnoses (Figure 2). The magnitude of the experts’ disagreement about some cases is even more obvious when their evaluations of probability are compared with cases that might be RA, SpA or UA. The range of variation was quite wide (0 to 91 mm on a visual analog scale for RA and 2 to 100 mm for SpA). There
was also a large overlap among the 3 possible diagnoses, especially when extreme choices were considered and not just the 50% of centiles close to the median (Figure 3).

Hence, our results suggest that there is still no general agreement about the nosology of early SpA and UA, as emphasized previously12-14, and that diagnoses could be partially influenced by cultural habits. For instance, the fact that French experts were more familiar than the international panel with the concept of SpA, but less familiar with that of UA, could be related to the particular interest of some French experts in SpA8,15.

Last but not least, the experts' choices were not strongly associated with satisfaction of ACR criteria for RA or ESSG criteria for SpA. This suggests that the conventional way of developing classification criteria (i.e. by trying to model consensus opinion) is inappropriate for the development of meaningful diagnostic criteria for early arthritis, and that other models for predicting the outcome of recent onset arthritis in nonresearch settings should be optimized.

Of course, this survey has obvious limitations that preclude dogmatic conclusions. First of all, the number of experts was limited. Second, it is likely that discrepancies would have been less striking if each expert had examined real patients rather than paper cases. Third, experts did not agree in advance about the description of cases. Fourth, they were not advised to use classification criteria. Moreover, 7/10 cases were selected out of 157 atypical cases during followup of a cohort of 270 patients with early arthritis because they were difficult to classify, indicating that they were not representative of the whole range of patients with early arthritis.

Nonetheless, our survey is a valuable illustration of the need to optimize current classification criteria and improve their efficiency within the context of early RA16 and early SpA. Although earlier reports reached more optimistic conclusions12, recent studies have questioned the efficiency of ACR 1987 criteria. In their very large population-based study on the outcome of early arthritis, Harrison et al noted that the percentage of early arthritis classified as RA could change from 38 (list format) to 67% (tree format) at baseline, and from 25 to 82% after 3 years of followup, depending on the way in which the 1987 ACR criteria for RA are applied (cumulatively or not, and using either the tree or the list format)16. This suggests a need for more precise definition of how the criteria should be added. Another possible obstacle to optimization of ACR criteria for RA is that early RA and early SpA overlap more frequently than is generally assumed. This is the impression given by some of the experts in our survey. A recent study also indicated that a significant proportion of patients with early arthritis tends to fulfil the criteria for both RA and SpA and share predisposing genetic factors for both disorders18. This is also apparent in the irregular nosology of psoriatic arthropathy, which can fulfil criteria for both RA and SpA (although this was not true for Case 10 in the present study). Such overlaps between early RA and SpA could easily have been ignored in previous reports. According to current ESSG criteria for SpA and 1987 criteria for RA, classifications as SpA or RA are frequently mutually exclusive, depending on whether arthritis is symmetrical or not. In fact, this criterion has been poorly defined: How should a patient with asymmetrical arthritis of the hands, but symmetrical arthritis of the ankles, be classified? Hence there could also be a need for more precise definitions and/or changes in current ACR and ESSG criteria.

Ideally, the question as to whether some disorders classified as RA or SpA share common pathogenic pathways (either environmental or genetic factors) should be resolved first. The solution to this difficult issue might indeed afford new diagnostic (and prognostic) tools.
www.cri-net.com

(including genetic markers, microbiological assays, and advanced imaging) that could reconcile experts about the classification of early arthritis and reduce the percentage of patients who remain unclassified. Indeed, as recently proposed by Schumacher and Bardin, the concepts of early RA and early SpA, still diagnosed mainly on clinical presentation, could be progressively replaced by new classifications based on a listing of the predisposing factors present (e.g. infectious triggers and features of host response). However, this approach would need to be validated by longterm followup of early arthritis cohorts. To achieve this goal, atypical early arthritis, such as the cases detailed in our survey, should not be excluded from clinical and experimental studies. In depth analyses of these frustrating cases at the clinical level may tell us even more about the pathogenesis of early and chronic rheumatism than do typical RA and SpA.

APPENDIX 1. Ten case descriptions used in the survey. MCP: metacarpophalangeal joint; PIP: proximal interphalangeal joint; CRP: C-reactive protein; ESR: erythrocyte sedimentation rate; IPD:

Case 1. A 39-year-old Caucasian male. Family history: mother diagnosed with RA; one brother with psoriasis. May 1996: knees painful and swollen. Morning stiffness = 30 min. Right big toe painful and swollen like a sausage. June 1996: right shoulder, left forefoot, wrists and hands (MCP and PIP) painful. Right shoulder, wrists, and knees swollen for 2 months. Morning stiffness = 45 min. Cervical spine painful at night and stiff in the morning. Loss of 15 kg within the past 6 months. No urogenital, digestive, or ocular signs. No personal dermatological signs. ESR = 92 mm; CRP = 68 mg/l. Latex + [200 IU/ml (normal range = 0-25)], Rose-Waaler + [32 IU/ml (normal range = 0-12)]. RF-IgM (ELISA): 0.651 (normal < 0.200). Antinuclear antibodies negative. HLA-B27 positive. HLA-DRB1 02 and 13, DQB1 0302-0609. September 1996: new flare of arthritis: swelling of wrists, MCP II and III, PIP II and III (left and right sides) and ankles. Forefeet and cervical spine also painful for 30 min in the morning. April 1997: 15 kg lost have been regained. Patient denies any pain (Lee = 3, Ritchie = 0). Latex and Rose-Waaler now negative, as well as RF-IgM. October 1997: cervical spine painful and stiff in the morning. Left wrist tender. March 1998: patient denies any symptoms (Lee = 1, Ritchie = 0). X-rays (pelvis, feet, hands) normal. Overall followup: 22 months.


Case 3. A 38-year-old Caucasian female. March 1996: slight diarrhea 1 month before first arthritis. Slight arthritis of wrists and shoulders for 2 months (swelling, morning stiffness < 45 min, and pain during mobilization). Forefeet also tender, and one knee swollen. ESR = 18 mm, CRP = 5.4 mg/l. Antinuclear antibodies negative. Latex, Rose-Waaler and RF-IgM

**Case 4.** A 52-year-old Caucasian female. No family or personal history of arthritis. July 1995: 3 weeks after a 7-day diarrhea (no stool culture), arthritis of right hip, left knee, ankles and PIP (both sides). Morning stiffness = 90 min. Antinuclear antibodies negative. Latex, Rose-Waaler, RF-IgM, RF-IgG, RF-IgA all negative. HLA-B27 negative. DRB1 01 and 13, DQB1 0604 and 0303. ESR = 36 mm; CRP = 3.2 mg/l. June 1996: arthritis of knees, arthralgia of wrists and fingers (PIP and MCP). November 1996: ongoing synovitis of right hip (leading to a prosthesis) and both ankles. Arthralgia of MTP III and IV (right and left). May 1998: synovitis of left hip and knee. Tenderness of sternoclavicular joints. Patient denies other present or past symptoms, and especially dermatological, urogenital or ocular signs. X-rays normal except for total chondrolysis of right hip (prosthesis). Overall followup: 34 months.

**Case 5.** A 51-year-old Caucasian male. No history of family or personal disorder, except slight personal bronchiectasis (CT-scan) and one flare of psoriasis 15 years before. December 1995: inflammatory pain in shoulders, knees and ankles. May 1996: synovitis of knees (large swelling). December 1996: inflammatory pain in ankles and shoulders. Morning stiffness: 2 h. Antinuclear antibodies positive (1:100, homogeneous). Latex positive [48 IU/ml (normal range < 25)], Rose-Waaler dubious [10 IU (normal range < 12)], RF-IgM ELISA positive (0.258; normal < 0.200), HLA- B27 negative, DRB1 0707, DQB1 0202 (homozygous for both). ESR = 28 mm; CRP = 15 mg/l. January 1998: synovitis of knees and ankles. MCP left IV and PIP right II tender on palpation. May 1998: Morning stiffness > 1 h (knees mainly). Tenderness of MCP left IV and III, right III and II. Slight arthritis of knees and ankles. Patient denies other present or past symptoms, and especially gastrointestinal, urogenital or ocular signs. X-rays (pelvis, feet, hands) normal. Overall followup: 28 months.


Case 8. A 37-year-old Caucasian male. Family history of ankylosing spondylitis (5 cases). November 1995: synovitis of the right knee, right wrist and left ankle for 3 months. Inflammatory spinal pain improved by exercise. Morning stiffness of joints: 90 min. ESR = 14 mm; CRP = 10 mg/l. Latex and RF-IgM (ELISA) negative. Anti-nuclear antibodies negative. HLA-B14 and B40. HLA-DRB1 04 and 15. November 1995 to October 1997: transient inflammatory spinal pain for 3 months. Patient denies other present or past symptoms, and especially gastrointestinal, urogenital or ocular signs. X-rays (pelvis, feet, hands) normal. Overall followup: 23 months.

Case 9. A 31-year-old Caucasian female. No family or personal history of arthritis. December 1995: synovitis of the right wrist, right PIP IV, right MCP IV, and left MCP II. Shoulders, right MTP IV and left MTP II painful. Morning stiffness = 2 h. ESR = 40 mm; CRP = 57 mg/l. Latex and RF-ELISA [0.162 (normal < 0.200)] negative. Antinuclear antibodies negative. HLA-B8 and B44. HLA-DRB1 04 and 07. July 1996: synovitis of the right wrist, right MCP IV, and right PIP V. All MTP painful. June 1998: synovitis of the right wrist. All MTP painful. Latex and RF-IgM still at the limit of positivity. Patient denies other present or past symptoms, and especially gastrointestinal, urogenital or ocular signs. X-rays: feet = erosion of both MTP V; hands = erosion and demineralisation of right wrist; pelvis = normal. Overall followup: 31 months.

Case 10. A 23-year-old Caucasian male. No family or personal history of arthritis. Father suffering from psoriasis. March 1996: synovitis of right PIP II-III, and left PIP for 2 months. Morning stiffness = 15 min. ESR = 2 mm; CRP = 3 mg/l. Latex and RF-IgM negative. Antinuclear antibodies negative. HLA-B8 and B16. HLA-DRB1 01 and 07. October 1996: flare of psoriasis, but no more arthritis. September 1997 to April 1998: pain alternating in both buttocks. Patient denies other present or past symptoms, and especially gastrointestinal, urogenital or ocular signs. X-rays (pelvis, feet, hands) = sacroiliitis. Overall followup: 26 months.

From the Department of Rheumatology, Nantes University Medical School, CHU Nantes, and the Department of Rheumatology, Brest University Medical School, CHU Brest, France.

Supported by the 1995 Clinical Research Hospital Program.

J-M. Berthelot, MD, Department of Rheumatology, Nantes University Medical School; M. Klarlund, MD, Department of Rheumatology, Hvidovre University Hospital, Denmark; D. McGonagle, MRCPI, Department of Rheumatology, Leeds Hospital, United Kingdom; H.J. Bernelot-Moens, MD, Medisch Spectrum Twente, Enschede, The Netherlands; A. Calin, MD, FRCP, Royal National Hospital for Rheumatic Diseases, Bath, United Kingdom; B.J. Harrison, MRCP, ARC Epidemiology Research Unit, University of Manchester, Manchester, United Kingdom; H.R. Schumacher Jr, MD, Department of Rheumatology, Hospital of the University of Pennsylvania, Philadelphia, PA, USA; K. Kaarela, MD, Rheumatism Foundation Hospital, Heinola, Finland; A.A. Drosos, MD, FACR, Associate Professor of Medicine/Rheumatology, Division of Rheumatology, Department of Internal Medicine, Ioannina Medical School, Ioannina, Greece; J.L. Hülsemann, MD, Division of Rheumatology, Department of Internal Medicine, School of Medicine, Hannover, Germany; W-H. Koh, MBBS, MRCP (UK), FAMS, Consultant, Tan Tock Seng Hospital, Singapore; Y-T. Konttinen, MD, PhD, Department of Anatomy, University of Helsinki, Helsinki, Finland; L. Punzi, MD, PhD, Associate Professor of Rheumatology, Division of Rheumatology, Department of Medical and Surgical Sciences, Padua, Italy; K. Tanimoto, MD, Professor of Medicine,
Department of Rheumatology, Saitama University, Saitama, Japan; H.J. Williams, MD, Department of Internal Medicine, School of Medicine, Salt Lake City, UT, USA; F. Wolfe, National Data Bank for Rheumatic Diseases and University of Kansas School of Medicine, Wichita, KS, USA; C-A. Zerbini, MD, Department of Rheumatology, Hospital Heliopolis, Sao Paulo, Brazil; A. Saraux, MD, PhD, Department of Rheumatology, Brest University Medical School, CHU, Brest, France.

Address reprint requests to Dr. J.M. Berthelot, Rheumatology Unit, Nantes University Medical School, Hôtel-Dieu, CHU Nantes, 44093, Nantes Cedex 01, France.


REFERENCES


