POINTS OF VIEW

# The challenge of using the rheumatoid arthritis diagnostic criteria in clinical practice

Salvatore Corrao · Christiano Argano · Luigi Calvo · Giovanni Pistone

Received: 16 November 2014/Accepted: 27 January 2015/Published online: 12 February 2015 © SIMI 2015

Abstract The new 2010 ACR/EULAR (American College Rheumatology/European League of Against Rheumatism) criteria of Rheumatoid Arthritis recently published, have been released to classify and identify patients with early RA who could benefit from early therapy. They recommend anti-citrullinated protein antibody (ACPA) testing as an alternative criterion to Rheumatoid Factor (RF) and ACPA that were introduced together with the other classic criteria in a scoring system. We previously criticized these new criteria because of unavailable specificity and sensibility in the first paper, and the use of ACPA as dichotomous criterion (presence/absent) and alternatives to rheumatoid factor. Our previous work promoted discussion and fostered new research on this issue. By the light of new data, in an effort to improve clinical reasoning, we suggest a more practical probabilistic point of view. In this regard, we analyze the sensitivity and specificity of the diagnostic studies that evaluate the performance of the 2010 classification criteria. Then, we compare the old and the new classification criteria. Subsequently, we describe the use of likelihood ratios applied to the classification criteria and different cutoff levels of ACPA for decisionmaking in different setting. Moreover, we define some properties of likelihood ratios and their use for diagnosing or excluding rheumatoid arthritis. We want to share this

S. Corrao · L. Calvo · G. Pistone

kind of knowledge within the scientific community because we believe that it can help general practitioners and specialists to recognize early arthritis patients implementing a more efficient probabilistic clinical reasoning.

**Keywords** Early rheumatoid arthritis · ACR/EULAR criteria · ACR criteria · Probabilistic clinical reasoning · Likelihood ratio · Sensitivity and specificity

#### Introduction

The early recognition of rheumatoid arthritis (RA) has become an essential issue in clinical practice for early treatment to improve patient outcomes, and to block disease progression by avoiding joint destruction and functional disability [1]. In the past decade much interest has been focused on anti-citrullinated peptide/protein antibodies (ACPA), claimed as a significant serological marker in the diagnosis of early RA and throughout the course of the disease [2].

For this reason, the new 2010 ACR/EULAR (European League Against Rheumatism) criteria have been specifically prepared to classify and identify patients with early RA who might benefit from early DMARD therapy introducing ACPA as alternative criterion to rheumatoid factor (RF). However, we previously underlined some methodological matters regarding these new criteria: (1) non-homogeneity of performance of the ROC curves among the three cohorts used for the formulation of the criteria, (2) absence of computation of specificity and sensitivity that prevented the comparison of performance with the previous 1987 criteria in the first paper of the criteria set, and (3) use of ACPA as a dichotomous criterion instead of considering it as a continuous variable with

S. Corrao  $(\boxtimes) \cdot C$ . Argano

Centre of Research for Effectiveness and Appropriateness in Medicine (CREAM); Biomedical Department of Internal Medicine and Subspecialties [DiBiMIS], University of Palermo, Piazza delle Cliniche n°2, 90127 Palermo, Italy e-mail: s.corrao@tiscali.it

National Relevance Hospital Trust, ARNAS Civico, Di Cristina e Benfratelli, Piazza Nicola Leotta nº4, 90127 Palermo, Italy

different diagnostic performance according to different cutoff values [3].

On the contrary, these criteria are a valuable effort by ACR/EULAR experts to classify patients for research purpose.

Several authors such as Zeidler and Montecucco highlight the need for a better classification and diagnosis of early arthritis because of the risk of misclassification and unjustified treatment [4–6].

The aim of this manuscript is to analyze the available literature data to give the best clinical decision-making tools for clinicians involved in RA diagnosis by the utilization of positive and negative Likelihood Ratios.

# The old and the new criteria

The American College of Rheumatology (ACR) 1987 RA criteria have been used for diagnostic purposes in clinical practice. However, these criteria are limited by low sensitivity and poor specificity for targeting early arthritis, and therefore will fail to identify the crucial population that hopefully would benefit from early therapeutic intervention. The new 2010 ACR/EULAR (European League Against Rheumatism) criteria have been prepared to classify and identify patients with early RA, who might benefit from early therapy [6].

Several previous studies have evaluated the performance of the 2010 classification criteria and 1987 criteria [7–10].

Table 1 shows different outcomes such as methotrexate and DMARD therapy with or without radiographic information, doctor's diagnosis, subjects with 1-year swollen joints, or diagnosis of RA at follow-up. Notwithstanding, the specificity of 1987 criteria was higher than 2010 criteria in contrast to the sensitivity. Computed positive and negative Likelihood Ratios do not indicate a sufficient probability to diagnose and to exclude a disease.

Moreover two recent systematic reviews and metaanalyses (Table 2) performed by Sakellariou and Radner, respectively, considered methotrexate and DMARD as reference, show high sensitivity and low specificity of 2010 criteria as well as little informative power for computed LR + and LR- [5, 11].

Therefore, more subjects may be marked as RA initially and diagnosed on follow-up with another disease. The 2010 criteria are less able to differentiate RA from other types of inflammatory arthritis, and will incorrectly diagnose RA in patients having other types of arthritis.

Thus, in addition to systematic reviews, even the most recent literature available cites the weakness of the new 2010 criteria in clinical practice.

In this regard, Zeidler and Montecucco clearly emphasize the significant risk of misclassification and unjustified treatment applying the new criteria [5]. In this scenario, while the most relevant scientific societies produce criteria, the physician remains alone in front of the patients, with the risk of misdiagnosis and the probability of administering inappropriate care.

How can a physician manage all this uncertainty arising from the lack of relevance and specificity of these criteria? A possible solution is to redirect clinical reasoning from excessive simplification, using old or new diagnostic criteria, managing uncertainty through a more comprehensive probabilistic approach. Such an outlook demands very hard work.

### The usefulness of ACPA

In past few years, the ACPA have emerged as a valuable serologic marker of RA, while rheumatoid factor had been the only biomarker considered in the past. Indeed, ACPA was included in the 2010 RA classification criteria [6]. However, we may specifically use the evidence about ACPA to discuss the diagnostic criteria according to a probabilistic approach. Nishimura and colleagues show a better diagnostic specificity of ACPA in comparison to Rheumatoid Factor, and according to their analysis, ACPA may also be a better predictor of erosive disease: the pooled sensitivity and pooled specificity were 67 and 95 %, respectively [12]. Moreover, a recent systematic review has been published on the accuracy of ACPA in diagnosing early RA. This analysis of 151 studies shows hat sensitivity and specificity are 57 and 96 %, respectively (analysis refers only to 15 relevant cohort studies) [13].

ACPA appears to be a reliable predictor of erosive RA, making it a potentially important prognostic tool that might be used to direct patient management decisions.

Recently Bizzarro and colleagues find that the presence of ACPA, at both low and high concentration, is significantly found in RA development in subjects with recent onset arthritis. Moreover, the time interval from the onset of the first symptoms to the fulfillment of the classification criteria seems to be directly related to the initial ACPA level [14].

Other authors have tried to test the performances of combining ACPA with the ACR 1987 criteria to diagnose early RA, without any improvement [15].

Conversely, Liao et al. [16] in a retrospective study, find that the sensitivity increases up to 63 % in patients with symptoms less than 6 months, and specificity decreases from 86 to 72 %. Moreover, data from the French ESPOIR cohort show that the addition of ACPA positivity to the enhanced 1987 criteria increases the number of patients correctly classified as having early stage RA [17].

References and study design	1987 criteria sensitivity	2010 criteria sensitivity	1987 criteria specificity	2010 criteria specificity	1987 LR+	2010 LR+	1987 LR-	2010 LR-
Berglin E et al. [22] prospective study	68 % (MTX within first year as outcome)	79 % (MTX within first year as outcome)	84 % (MTX within first year as outcome)	54 % (MTX within first year as outcome)	4.25	8	0.38	0.31
	72 % (diagnosis of RA at one- year-follow-up)	83 % (diagnosis of RA at one- year-follow-up)	91 % (diagnosis of RA at one- year-follow-up)	65 % (diagnosis of RA at one- year-follow-up)				
Biliavska I et al. [23]	55 % (DMARD start)	80 % (DMARD start)	76 % (DMARD start)	63 % (DMARD start)	2.29	2.16	0.59	0.32
double blind placebo	65 % (Diagnosis: RA)	85 % (Diagnosis: RA)	80 % (Diagnosis: RA)	64 % (Diagnosis: RA)	3.25	2.36	0.44	0.23
	57 % (Diagnosis: RA and/or DMARD start)	79 % (Diagnosis: RA and/or DMARD start)	0.80 (Diagnosis: RA and/or DMARD start)	66 % (Diagnosis: RA and/or DMARD start)	2.85	2.32	0.54	0.32
Makinen H et al. [10] prospective study	70 % (Erosive disease at 10 years)	87 % (Erosive disease at 10 years)	47 % (Erosive disease at 10 years)	44 % (Erosive disease at 10 years)	1.32	1.5	0.64	0.3
Kennish L et al. [24] prospective study	93 %(Doctor's diagnosis as gold standard)	97 % (Doctor's diagnosis as gold standard)	76 % (Doctor's diagnosis as gold standard)	55 % (Doctor's diagnosis as gold standard)	3.88	2.16	0.09	0.06
Britsemmer K et al. [8] Prospective study	76 % (MTX within one year as gold standard)	85 % (MTX within one year as gold standard)	59 % (MTX within one year as gold standard)	50 % (MTX within one year as gold standard)	1.85	1.7	0.41	0.3
	81 % (Expert opinion as gold standard)	90 % (Expert opinion as gold standard)	59 % (Expert opinion as gold standard)	48 % (Expert opinion as gold standard)	1.98	1.73	0.32	0.21
Kaneko Y et al. [9] Retrospective single centre observational study	47.1 % (subjects with swollen joints not better explain by other diseases)	73.5 % (subjects with swollen joints not better explain by other diseases)	92.9 % (subjects with swollen joints not better explain by other diseases)	71.4 % (subjects with swollen joints not better explain by other diseases)	6.63	2.57	0.57	0.37
van der Linden MP et al. [7] Prospective study	. 61 % (MTX therapy as outcome and without radiographic information)	84 % (MTX therapy as outcome and without radiographic information)	74 % (MTX therapy as outcome and without radiographic information)	60 % (MTX therapy as outcome and without radiographic information)	2.35	2.1	0.53	0.27
	61 % (MTX therapy as outcome and with radiographic information)	88 % (MTX therapy as outcome and with radiographic information)	74 % (MTX therapy as outcome and with radiographic information)	54 % (MTX therapy as outcome and with radiographic information)	2.35	1.91	0.53	0.22
	54 % DMARD initiation as outcome without radiographic information)	74 % DMARD initiation as outcome without radiographic information)	87 % DMARD initiation as outcome without radiographic information)	74 % DMARD initiation as outcome without radiographic information)	4.15	2.85	0.53	0.35
	54 % DMARD initiation as outcome with radiographic information)	79 % DMARD initiation as outcome with radiographic information)	87 % DMARD initiation as outcome with radiographic information)	68 % DMARD initiation as outcome with radiographic information)	4.15	2.47	0.53	0.31
	53 % (Persistent arthritis over five years as outcome without radiographic information)	71 % (Persistent arthritis over five years as outcome without radiographic information)	75 % (Persistent arthritis over five years as outcome without radiographic information)	65 % (Persistent arthritis over five years as outcome without radiographic information)	2.12	2.03	0.63	0.45
	53 % (Persistent arthrithis over five years as outcome with radiographic information)	77 % (Persistent arthrithis over five years as outcome with radiographic information)	75 % (Persistent arthrithis over five years as outcome with radiographic information)	56 % (Persistent arthrithis over five years as outcome with radiographic information)	2.12	1.75	0.63	0.41



References	Study design	2010 criteria sensitivity	2010 criteria specificity	LR+	LR-				
Sakellariou et al. [5]	Systematic review and meta-analysis	80 % methotrexate as standard	61 % methotrexate as standard	2.1	0.33				
		73 % DMARDs as standard	74 % DMARDs as standard	2.81	0.37				
Radner H et al. [11]	Systematic review and meta-analysis	85 % methotrexate as standard	52 % methotrexate as standard	1.77	0.29				
		80 % DMARDs as standard	65 % DMARDs as standard	2.29	0.31				

 Table 2
 List of systematic reviews, their results of diagnostic studies about the ACR/EULAR 2010 criteria and computed positive and negative likelihood ratios

Positive likelihood ratio sensitivity/(1-specificity), Negative likelihood ratio (1-sensitivity)/specificity

# The diagnosis of RA: beyond diagnostic criteria

All these data are likely to create confusion in the clinician whose only objective is to improve efficient clinical practice. Accordingly, Baye's theorem could help the physician to resolve this issue. In fact, Baye's theorem attests that the Likelihood Ratio of positive (LR+) and negative (LR-) results enables one to compute posterior probability using different "a priori" probabilities, in contrast with positive or negative predictive values. Specifically, Likelihood Ratios are sufficiently stable characteristics of a test having more informative power compared to sensitivity and specificity that are usually confusing for the clinician. In this regard, sensitivity and specificity are useful parameters when their values are near 100 % (a negative result of a test with sensitivity of 100 % enables the exclusion of a disease, while a test with specificity of 100 % is highly indicative of disease). This is not the case for the 2010 diagnostic criteria and ACPA. Therefore, how can we apply a sensitivity and specificity that do not reach the value of 100 %? Likelihood Ratios of a positive and negative test are always informative and easily usable as a counterpoint to sensitivity and specificity. In general, a positive Likelihood Ratio  $\geq 15$  and a negative Likelihood Ratio <0.15, respectively, indicate a good probability of diagnosing a disease, and a good probability of excluding a disease [18]. Moreover, Likelihood Ratios have another absorbingly interesting characteristic: their values can be multiplied between or among Likelihood Ratios for various tests. In our case, ACPA and the old American College of Rheumatology criteria can be considered to be two independent tests for the disease. Indeed the old criteria do not contain the ACPA test. For this reason, we can multiply the LR+ by themselves. The 1987 ACR criteria have an LR+ of 8.54 and an LR- of 0.07 (computed from the original sensitivity and specificity values of the American College of Rheumatology criteria). Data calculated from the systematic review of Within indicate that ACPA has a LR+ of 12.7 and a LR- of 0.45 [13]. It is clear that neither positive Likelihood Ratio nor negative Likelihood Ratio is easily used in clinical practice for diagnosing or excluding Rheumatoid arthritis. However, if we multiply their values by themselves  $(LR + = 8.54 \times 12.7 = 108.4; LR - =$  $0.07 \times 0.45 = 0.031$ ) the subsequent Likelihood Ratios (positive and negative) became highly informative even for an inexperienced physician. However, this would simply and correctly classify only subgroups of patients who are positive or negative for both the criteria. Nevertheless, these LR values remain highly informative and usable in different clinical practice settings (e.g. general practitioner or rheumatologic outpatient clinic). In particular, low pre-test probabilities need high positive likelihood ratios (to confirm a diagnosis), or, on the contrary, pretty low negative likelihood ratios (to exclude a disease). This is the case of pre-test probability of various diseases in the general population including rheumatoid arthritis. Thus, in the above-mentioned situations, a frequent use of Baye's theorem will generate a high post-test probability. Otherwise, an experienced clinician such as a rheumatologist may use his expertise to assign probabilities beyond a probabilistic logic so that the criteria might be unneeded by being accurately overcome by physician gestalt perception of pre-test probability [19]. In this case, a single lab test with optimal likelihood ratios, like ACPA, might be used to transform pre-test probability excluding or confirming the diagnosis. Obviously, the correct diagnosis is of paramount importance for an effective decision-making about specific drug prescriptions [20].

In this way, the use of combined information simplifies diagnostic utilization of ACPA cutoff values. Recently, Pietrapertosa et al. [21] describe the importance of different cutoff values to modify Likelihood Ratios according to Sackett lessons [18]. In fact, different cutoff values can yield the so-called SpIn (Specificity in: high values for the highest specificity, consequently to diagnose disease), and SNout (Sensitivity out: low values for the highest sensitivity, consequently to exclude disease). Indeed, high ACPA cutoff values may again overcome the use of the criteria. APCA values  $\geq 15.0$  U/mL correspond to positive Likelihood Ratios of 42 (values  $\geq 30.0$  U/mL actually correspond to positive LH =  $\infty$ ) becoming very useful in diagnosing early RA much more precociously, and without the use of the old and new criteria.

### Conclusion

Our analysis might be useful in enhancing the body of knowledge to improve clinical reasoning approach uncertainty in different settings. In particular, we discuss how the in-depth development of probabilistic logic applied to this notable topic using literature data allows one to cope with the uncertainty from different points of view. A general practitioner could use the new criteria only to exclude rheumatoid arthritis, while the fulfillment of the criteria needs a rheumatologic evaluation to possibly confirm the diagnosis. On the other hand, the rheumatologist might use a more comprehensive gestalt perception approach. In conclusion, we believe that this is the time to share this kind of complex reasoning within the scientific community to improve clinical practice.

**Conflict of interest** All the authors declare that they have no conflicts of interest.

#### References

- Van der Helm-van Mil AH, le Cessie S, van Dongen H, Breedveld FC, Toes RE, Huizinga TW (2007) A prediction rule for disease outcome in patients with recent-onset undifferentiated arthritis. How to guide individual treatment decisions. Arthritis Rheum 56:433–440
- Wiik AS, van Venrooij WJ, Pruijn GJ (2010) All you wanted to know about anti-CCP but were afraid to ask. Autoimmun Rev 10:90–93
- Corrao S, Calvo L, Licata G (2011) The new criteria for classification of rheumatoid arthritis: what we need to know for clinical practice. Eur J Intern Med 22(3):217–219
- Zeidler H (2013) Systemic literature review of the performance of the 2010 ACR/EULAR classification criteria for rheumatoid arthritis: good news of debatable significance. Ann Rheum Dis 72(8):e21
- Sakellariou G, Scirè CA, Zambon A, Caporali R, Montecucco C (2013) Performance of the 2010 classification criteria for rheumatoid arthritis: a systematic literature review and a metaanalysis. PLoS One 8(2):e56528
- Aletaha D, Neogi T, Silman AJ, Funovits J, Felson DT, Bingham CO 3rd et al (2010) 2010 rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. Arthritis Rheum 62:2569–2581
- van der Linden MP (2011) Knevel R, Huizinga TW, van der Helmvan Mil AH. Classification of rheumatoid arthritis: comparison of the 1987 American College of Rheumatology criteria and the 2010 American College of Rheumatology/European League Against Rheumatism criteria. Arthritis Rheum 63(1):37–42
- Britsemmer K, Ursum J, Gerritsen M, van Tuyl LH, van Schaardenburg D (2011) Validation of the 2010 ACR/EULAR classification criteria for rheumatoid arthritis: slight improvement over the 1987 ACR criteria. Ann Rheum Dis 70(8):1468–1470
- Kaneko Y, Kuwana M, Kameda H, Takeuchi T (2011) Sensitivity and specificity of 2010 rheumatoid arthritis classification criteria. Rheumatology 50(7):1268–1274

- Mäkinen H, Kaarela K, Huhtala H, Hannonen PJ, Korpela M, Sokka T (2013) Do the 2010 ACR/EULAR or ACR 1987 classification criteria predict erosive disease in early arthritis? Ann Rheum Dis 72:745–747
- Radner H, Neogi T, Smolen JS, Aletaha D (2014) Performance of the 2010 ACR/EULAR classification criteria for rheumatoid arthritis: a systematic literature review. Ann Rheum Dis 73(1):114–123
- Nishimura K, Sugiyama D, Kogata Y, Tsuji G, Nakazawa T, Kawano S et al (2007) Meta-analysis: diagnostic accuracy of anticyclic citrullinated peptide antibody and rheumatoid factor for rheumatoid arthritis. Ann Intern Med 146:797–808
- Whiting PF, Smidt N, Sterne JA, Harbord R, Burton A, Burke M et al (2010) Systematic review: accuracy of anticitrullinated peptide antibodies for diagnosing rheumatoid arthritis. Ann Intern Med 152:456–464
- 14. Bizzaro N, Bartoloni E, Morozzi G, Manganelli S, Riccieri V, Sabatini P et al (2013) Anti-cyclic citrullinated peptide antibody titer predicts time to rheumatoid arthritis onset in patients with undifferentiated arthritis: results from a 2-year prospective study. Arthritis Res Ther 15:R16
- 15. Le Loët X, Strotz V, Lequerré T, Boumier P, Pouplin S, Mejjad O et al (2011) Combining anti-cyclic citrullinated peptide with the American College of Rheumatology 1987 criteria failed to improve early rheumatoid arthritis diagnosis in the communitybased very early arthritis cohort. Rheumatology (Oxford) 50(10):1901–1907
- Liao KP, Batra KL, Chibnik L, Schur PH, Costenbader KH (2008) Anti-cyclic citrullinated peptide revised criteria for the classification of rheumatoid arthritis. Ann Rheum Dis 67(11):1557–1561
- 17. Nicaise-Roland P, Nogueira L, Demattei C, de Chaisemartin L, Rincheval N, Cornillet M et al (2013) Autoantibodies to citrullinated fibrinogen compared with anti-MCV and anti-CCP2 antibodies in diagnosing rheumatoid arthritis at an early stage: data from the French ESPOIR cohort. Ann Rheum Dis 72(3):357–362
- Sackett DL, Haynes RB, Guyatt GH, Tugwell P (1991) Clinical Epidemiology: a basic science for clinical medicine. Little Brown, Boston
- Trowbridge RL, Rutkowski NK, Shojania KG (2003) The rational clinical examination. Does this patient have acute cholecystitis? JAMA 289(1):80–86
- 20. Gualtierotti R, Casella F, GrAM (Gruppo di Autoformazione Metodologica) (2014) Is it safe to withdraw etanercept in established rheumatoid arthritis after low disease activity achievement? Intern Emerg Med 9(2):223–224
- Pietrapertosa D, Tolusso B, Gremese E et al (2010) Diagnostic performance of anti-citrullinated peptide antibodies for the diagnosis of rheumatoid arthritis: the relevance of likelihood ratios. Clin Chem Lab Med 48(6):829–834
- 22. Berglin E, Dahlqvist S (2013) Comparison of the 1987 ACR and 2010 ACR/EULAR classification criteria for rheumatoid arthritis in clinical practice: a prospective cohort study. Scand J Rheumatol 42(5):362–368
- 23. Biliavska I, Stamm TA, Martinez-Avila J, Huizinga TW, Landewé RB, Steiner G et al (2013) Application of the 2010 ACR/ EULAR classification criteria in patients with very early inflammatory arthritis: analysis of sensitivity, specificity and predictive values in the SAVE study color. Ann Rheum Dis 72(8):1335–1341
- Kennish L, Labitigan M, Budoff S, Filopoulos MT, McCracken WA, Swearingen CJ, Yazici Y (2012) Utility of the new rheumatoid arthritis 2010 ACR/EULAR classification criteria in routine clinical care. BMJ Open 2(5):e001117. doi:10.1136/ bmjopen-2012-001117