RHEUMATOLOGY

Rheumatology 2020;0:1–8
doi:10.1093/rheumatology/keaa061

Original article

Diagnosis of fibromyalgia: comparison of the 2011/2016 ACR and AAPT criteria and validation of the modified Fibromyalgia Assessment Status

Fausto Salaffi[1], Marco Di Carlo [1], Sonia Farah[1], Fabiola Atzeni[2], Dan Buskila[3], Jacob N. Ablin[4], Winfried Häuser[5] and Piercarlo Sarzi-Puttini[6]

Abstract

Objective. To compare the concordance of the three diagnostic criteria, respectively the 2011 ACR criteria (ACR 2011 Cr), the ACR 2016 criteria (ACR 2016 Cr) and the Analgesic, Anesthetic, and Addiction Clinical Trial Translations Innovations Opportunities and Networks (ACTTION)-APS Pain Taxonomy criteria (AAPT Cr), and to explore the performance of an additional set of criteria, the modified Fibromyalgia Assessment Status (FAS 2019 modCr), in the diagnosis of FM syndrome.

Methods. Consecutive patients with chronic widespread pain, referred by the primary care setting, underwent rheumatologic assessment that established the presence or not of FM and were investigated through the four sets of proposed criteria. For the FAS 2019 modCr, discriminant validity to distinguish patients with FM and non-FM was assessed with receiver operating characteristic curve analysis.

Results. A total of 732 (405 with FM and 327 with other common chronic pain problems) patients were evaluated. Against the clinical diagnosis of FM, the sensitivity, specificity and correct classification were, respectively: 79.8, 91.7 and 85.1% for ACR 2011 Cr; 78, 90.5 and 83.6% for the ACR 2016 Cr; and 73.8, 91.7 and 81.8% for the AAPT Cr. The alternative set, proposed on the FAS 2019 modCr, provided a maximal diagnostic accuracy with a score $\geq$20 (Youden's index), with a sensitivity of 84.2%, specificity 89.0% and positive likelihood ratio 7.65.

Conclusion. There is a considerable agreement between criteria-based diagnoses of FM, although the AAPT Cr perform least well in terms of percentage of correct classification. The FAS 2019 modCr had comparable characteristics.

Key words: fibromyalgia, chronic widespread pain, diagnosis, classification criteria

Rheumatology key messages

- The ACR 2011 criteria offer the best concordance with the clinical judgment.
- The AAPT criteria are those with the worst performance.
- The modified Fibromyalgia Assessment Status is valid for classification/diagnostic purposes.

Introduction

Fibromyalgia (FM) is a complex chronic pain condition that affects at least 2% of the adult population in western countries [1]. The uncertainties about this condition are still numerous, both regarding the pathophysiological mechanisms and the diagnostic/classification criteria [2]. Diagnosing FM is still challenging, taking an average of 2.3 years after first complaints [3].

Over the years, various classification, diagnostic and screening criteria have been developed. However there is no gold standard for the diagnosis of FM up to now.

[1]Rheumatological Clinic, Ospedale "Carlo Urbani", Università Politecnica delle Marche, Jesi (Ancona), [2]Rheumatology Unit, University of Messina, Messina, Italy, [3]Department of Medicine, H. Soroka Medical Center, Ben Gurion University of the Negev, Beer Sheva, [4]Department of Internal Medicine H, Tel Aviv Sourasky Medical Center, Tel Aviv, Israel, [5]Department of Internal Medicine 1, Klinikum Saarbrücken, Saarbrücken, Germany and [6]Rheumatology Unit, Internal Medicine Department, ASST Fatebenefratelli-Sacco, Milano Statale University School of Medicine, Milan, Italy

Submitted 7 October 2019; accepted 25 January 2020

Correspondence to: Marco Di Carlo, Rheumatological Clinic, Ospedale "Carlo Urbani", Università Politecnica delle Marche, Jesi (Ancona), Italy. E-mail: dica.marco@yahoo.it

To date, no work has compared the latest sets of criteria available.

Classification criteria are standardized definitions that are primarily intended to create well-defined, relatively homogeneous cohorts of patients for research purposes. Classification criteria are not intended to capture the entire world of possible patients, but rather to catch the majority of patients who share the key attributes of the condition. On the other hand, diagnostic criteria are a set of signs, symptoms and tests developed to be used in everyday clinical practice to lead the care of the individual. The diagnostic criteria are generally comprehensive and must reflect the different features (heterogeneity) of a disorder, with the aim of carefully identifying as many individuals with the disease as possible [4].

In 1990, the American College of Rheumatology (ACR) approved for the first time criteria for FM [5]. These classification criteria, which have been widely used in clinical practice, require the existence of chronic widespread pain (CWP) for >3 months, and evokable tenderness on digital palpation in at least 11 out of 18 specified tender points. The 1990 criteria were widely accepted by the research community, but generally ignored by most physicians, mainly because of the difficulty for non-rheumatologists in performing the tender point examination. Publication of the ACR preliminary diagnostic criteria for FM in 2010 (ACR 2010 Cr) eliminated the criterion of tender point examination [6]. This greatly facilitated the diagnostic process by eliminating the uncertainty deriving from the subjectivity of the tender point examination. The ACR 2010 Cr stated that FM can be defined as CWP associated with somatic symptoms, and recommended the use of the widespread pain index (WPI) and the symptom severity scale (SSS) as diagnostic measures. To simplify the ACR 2010 diagnostic process further, eliminating even the need for an interviewer allowing epidemiologic studies, a modification of these criteria was made in 2011 (ACR 2011 Cr). This set of criteria confirmed the utilization of the WPI and SSS, introducing the FM symptom scale, which is the sum of the WPI and SSS. The WPI comprises 19 areas of the body and the patient has to show where he or she had pain in the past week, scoring one point for each painful area (final score 0–19). The SSS is determined considering fatigue, unrefreshing sleep, cognitive manifestations and somatic symptoms. Each symptom is assigned a score between 0 and 3, according to its severity (in the case of the first three) or the amount (in case of somatic symptoms). It was considered that a patient met the diagnostic criteria with the presence of a WPI ≥7 and an SSS ≥5, or WPI 3–6 and SSS ≥9. It was found that an FM symptom scale ≥13 (out of a possible 31) provided a specificity of 91.8% and a sensitivity of 96.6% for a diagnosis of FM [7]. The FM symptom scale has been renamed the polysymptomatic distress scale (PDS). The term PDS will be used in this study [8]. In 2016, ACR further revised the FM diagnostic criteria (ACR 2016 Cr) by adding the generalized pain criterion, which was defined as pain in at least four of five regions (excluding jaw, chest and abdominal pain), specifying somatic symptoms as headache, pain or cramps in lower abdomen, and depression, and confirming the framework of PDS. These criteria state specifically that 'a diagnosis of FM is valid irrespective of other diagnoses. A diagnosis of FM does not exclude the presence of other clinically important illnesses.' This comment was inserted into the ACR 2016 Cr to make clear that FM criteria were valid in the presence of other clinically important illnesses, including those diagnosed using the ACR 2010 and 2011 criteria [9].

Recently, the Analgesic, Anesthetic, and Addiction Clinical Trial Translations Innovations Opportunities and Networks (ACTTION)- American Pain Society (APS) Pain Taxonomy criteria (AAPT) established an international FM working group, consisting of clinicians and researchers with expertise in FM, and introduced new criteria focused on (i) multisite pain, defined as six or more pain sites from a total of nine possible sites, (ii) moderate to severe sleep problems or fatigue, and (iii) multisite pain plus fatigue or sleep problems must have been present for at least 3 months [10].

These studies report a fairly wide variation in the sensitivity, specificity, diagnostic standards and spectrum of pain disorders that were sampled.

The aims of the current study were to compare the concordance of the three most recent and internationally accepted sets of diagnostic criteria (ACR 2011 Cr, ACR 2016 Cr and AAPT Cr) in the clinical setting and, in addition, to explore the performance of an alternative set of classification criteria developed using the modified Fibromyalgia Assessment Status (FAS) questionnaire. This new set of criteria has been called the FAS 2019 modified Criteria (FAS 2019 modCr).

Methods

Subject recruitment

The dominating principle of this study was to incorporate a wide range of common pain disorders as seen in everyday clinical practice. Adult patients ≥18 years old were enrolled from the practices of two rheumatologists (F.S. and M.D.C., clinicians with experience in the diagnosis and treatment of FM) with locations in the Rheumatological Clinic of the Università Politecnica delle Marche (tertiary centre), Jesi (Ancona), Italy, from June 2017 to May 2019. The subjects were referred by primary care practitioners to a rheumatologist for a history of CWP, defined as axial plus upper and lower body plus left- and right-sided pain [5]. The study complied with the Declaration of Helsinki, the research protocol was approved by the local ethics committee (DAV2 n. 1654) and informed consent has been obtained from the subjects.

Procedures

Patients suffering from FM were diagnosed on clinical grounds by two rheumatologists (F.S. and M.D.C.) who

trained together [11]. The decision regarding whether the patient could be clinically diagnosed as having FM for this study was made considering the long-term patient-clinician experience and included factors related to pain, tenderness, fatigue, sleep disturbance, comorbidity and psychosocial variables. In all subjects, the presence of FM was evaluated through the ACR 2011 Cr, ACR 2016 Cr and AAPT Cr. The rheumatologists' diagnosis of FM as the gold standard was conducted as per previous studies [5, 6], according to each rheumatologist's individual methodology. It was not a requirement of diagnosis to satisfy the ACR classification criteria [5]. The diagnosis of FM was made regardless of any other diagnosis. Therefore, no distinction was made between primary and secondary FM. In not making this distinction, we followed the recommendation in the report of the ACR FM criteria committee stating 'primary and secondary-concomitant FM were essentially indistinguishable with the study variables, and the criteria proposed worked equally well in both groups' [9]. Control subjects were patients with non-inflammatory painful disorders such as degenerative neck and back pain syndromes or focal myofascial pain, osteoarthritis (hip, knees or hands), tendonitis, chronic migraine headaches other headache disorder, or peripheral neuropathy who had not been diagnosed previously as having FM and who were of the same sex and were no more than 10 years younger or 10 years older than the FM case. Patients with any inflammatory rheumatic disorder (e.g. rheumatoid arthritis or psoriatic arthritis), active cancer, fractures, or other non-rheumatic causes of pain were excluded from the study. All eligible cases were screened by laboratory evaluation for the exclusion of other differential diagnoses. Accordingly, we excluded those with abnormal hemogram (haemoglobin <12 g/dl in females and <14 g/dl in males), elevated erythrocyte sedimentation rate (>15 mm/h in males and >20 mm/h in females) or high C-reactive protein level (>5 mg/l), or if there was any abnormality in thyroid function tests, serum calcium, phosphorus, alkaline phosphatase, blood urea nitrogen, creatinine, ferritin and creatine kinase. No attempt was made to determine interrater reliability for the diagnosis of FM among the two rheumatologists.

Development of alternative Fibromyalgia Assessment Status 2019 modified Criteria (FAS 2019 modCr)

The FAS 2019 modCr are the updated version of the FAS questionnaire. The validation of the original FAS was published in 2009 [12]. FAS includes questions addressing fatigue, quality of sleep and CWP. In FAS, patients are asked to rate CWP in 19 body regions; for each region pain is rated on a four-point scale (0 = no pain, 1 = mild pain, 2 = moderate pain, 3 = severe pain). Next to the CWP assessment, FAS assesses fatigue and sleep disturbances (each of which is scored by a numerical rating format from 0 to 10). In order to improve patient understanding and feasibility of FAS, and to reduce system variability, we decided to simplify the scoring system by considering only the presence (score 1) or absence (score 0) of pain in the body regions of a manikin, leaving out assessment on the 4-point numerical scales. This work includes these modifications of the FAS. This simplifying approach of a clinical instrument has already been described in the literature [13]. The final score of the FAS 2019 modCr ranges from 0 to 39 (Fig. 1).

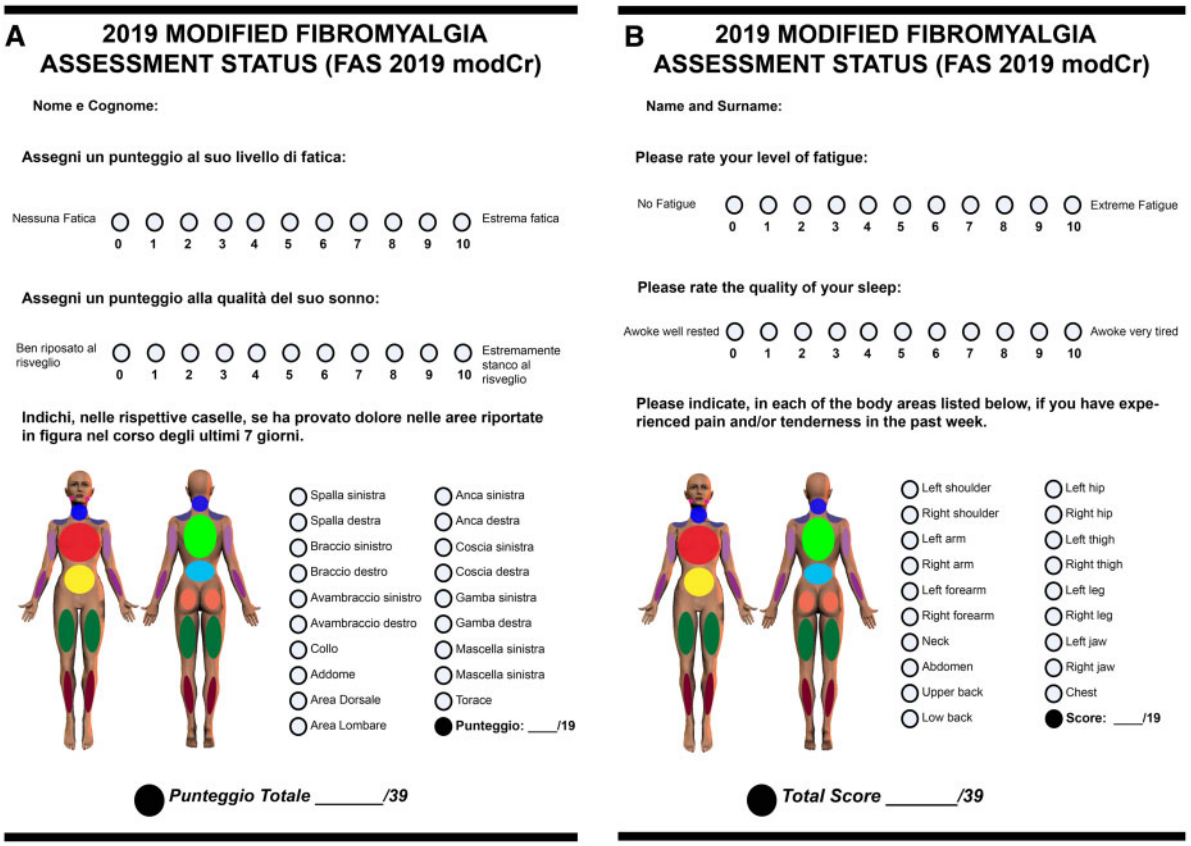
Questionnaire data

A comprehensive package of questionnaires including demographic data, disease duration, disease-related variables and quality of life items was administered to the patients. The demographic variables were: age, sex, marital status (single, married, divorced/separated) and level of education (primary, secondary, high school/university). FM severity was assessed through the Italian validated version of the revised Fibromyalgia Impact Questionnaire (FIQR) [14], which is the updated version of the Fibromyalgia Impact Questionnaire (FIQ) [15, 16]. Designed to override the limitations of the FIQ, FIQR is made up of 21 numerical rating scales (range 0–10, with 10 being the 'worst'), and investigates three main domains of health: function, overall impact and symptoms [17]. FIQR tries to improve the original scale by adding new questions related to memory, tenderness, balance and environmental sensitivity. The questions refer to the previous 7 days. The final score (range 0–100, with greater values indicating a worse severity) is the sum of the ratings of the three domains: the algebraic sum of the 9-items function domain (range 0–90) is divided by three, the algebraic sum of the 2-items overall impact domain (range 0–20) remains as it is, and the algebraic sum of the 10-items symptom domain (range 0–100) is divided by two. The FIQR and cut-off points for disease severity were: remission ≤ 30, mild severity >30 and ≤ 45, moderate severity >46 and ≤ 65, and high severity >65 [18].

Statistical analysis

Differences among groups in demographic and clinical characteristics were calculated with the unpaired *t*-test. If data were not sampled from Gaussian distributions, a nonparametric test (Mann–Whitney *U*-test) was used. Statistical significance was set at $P < 0.05$. Data were processed with the MedCalc Statistical Software, version 19.0 (Ostend, Belgium), for Windows XP. To compare categorical data, we used Fisher's exact test. Specificity and sensitivity of the ACR 2011 Cr, ACR 2016 Cr and AAPT Cr for the diagnosis of FM were determined using the rheumatologists' *a priori* diagnosis of FM as the gold standard, and determining the percentages of subjects who met the criteria for a diagnosis of FM set out in the different comparative criteria. We estimated diagnostic test characteristics using an on-line diagnostic test calculator (http://araw.mede.uic.edu/cgi-bin/testcalc.pl). In addition, we used the receiver operating characteristic (ROC) curve analysis to explore the discriminative accuracy of the FAS 2019 modCr, as

Fig. 1 The FAS 2019 modCr



(**A**) Italian and (**B**) English versions of the FAS 2019 modCr. The FAS 2019 modCr is a patient administered questionnaire comprising two sections. The first section contains two questions on symptoms of fatigue and unrefreshing sleep during the past week, each of which is scored by a numerical rating format from 0 (no problem) to 10 (severe problems). The scores are summed with a maximum score of 20. The second section comprises a regional pain scale assessing 19 areas of the body, on which the patient indicates where he or she had pain in the past week. The number of separate pain sites are summed from a maximum of 19 body sites (score 0–19). The final score of the FAS 2019 modCr ranges from 0 to 39. FAS 2019 modCr: Fibromyalgia Assessment Status 2019 modified Criteria.

compared with the PDS, to distinguish patients with FM and non-FM. Since ROC analysis requires external criteria to be dichotomous, clinical criteria were employed as external criteria. ROC curves were created by plotting the true-positive proportion (sensitivity) *vs* the false-positive proportion (100-specificity) for the discrimination between FM and non-FM patients for multiple cut-off points. From the ROC curves, we computed the optimal cut-off point corresponding to the maximum sum of sensitivity and specificity. The area under the ROC curve was calculated to quantify the discriminative accuracy. Areas under the ROC curve from 0.50 to about 0.70 represents poor accuracy, those from 0.70 and 0.90 are 'useful for some purposes', while higher values represent an optimal accuracy [19]. The non-parametric Wilcoxon's signed ranks test is used for calculation and comparison of the areas under the ROC curves, as suggested by Hanley and McNeil [20].

Results

Subject characteristics

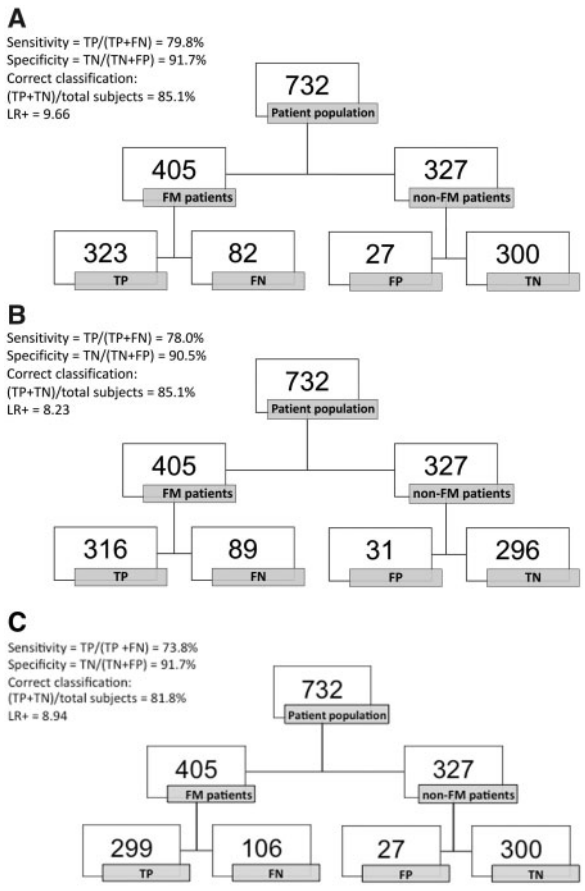
Complete data on 732 patients were evaluated, with 405 suffering from FM (94.1% females) and 327 (89.3% females) with other common chronic pain problems, such as degenerative neck (27 patients; 8.6%), back pain syndromes (36 patients; 11.0%), focal myofascial pain (20 patients; 6.1%), osteoarthritis (hip, knees or hands) (164 patients; 50.1%), tendonitis (28 patients; 8.5%), chronic migraine headaches or other headache disorder (41 patients, 12.4%), and peripheral neuropathy (11 patients; 3.3%). Sample demographic characteristics for 732 patients and subgroups were shown in Table 1.

Sensitivity, specificity, correct classification and positive likelihood ratio (LR+) of the ACR 2011 Cr, ACR 2016 Cr and AAPT Cr were determined *vs* the clinical diagnosis of FM.

Table 1 Demographic characteristics of the whole sample

Sample demographics	All patients (n = 732)	FM group (n = 405)	Non-FM group (n = 327)
Age [mean (s.d.)], years	52.08 (9.70)	52.45 (9.01)	51.60 (10.49)
BMI	26.21 (2.68)	26.38 (2.86)	25.99 (2.43)
Disease duration [mean (s.d.)], years	4.95 (4.56)	4.88 (4.22)	5.04 (4.95)
Marital status [number (percent married)]	378 (51.6)	191 (47.2)	187 (57.1)
Education level [number (percent)]			
Primary school	316 (43.16)	164 (40.49)	152 (46.48)
Secondary school	323 (44.12)	175 (43.20)	148 (45.25)
High school/university	93 (12.70)	49 (12.09)	44 (13.45)
Females [number (percent females)]	673 (91.9)	381 (94.1)	292 (89.3)
Males [number (percent males)]	59 (8.1)	24 (5.9)	35 (10.7)

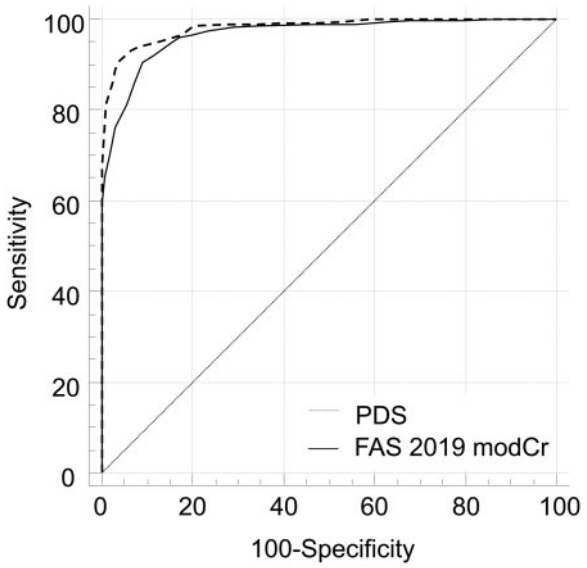
Fig. 2 Comparison of the three sets of diagnostic criteria *vs* the clinical diagnosis



Comparison of (**A**) the ACR 2011 criteria, (**B**) the ACR 2016 criteria and (**C**) the AAPT criteria *vs* the clinical criterion. FM: fibromyalgia; TP: true positive; TN: true negative; FP: false positive; FN: false negative; LR+: positive likelihood ratio.

The comparison of the sets of criteria *vs* the clinical criterion is shown in Fig. 2, with true positivity, false positivity, false negativity and true negativity. Applying the clinical diagnosis of FM as external criterion, the sensitivity, specificity, correct classification and LR+, respectively, resulted: 79.8%, 91.7%, 85.1% and 9.66 for ACR 2011 Cr; 78%, 90.5%, 83.6% and 8.23 for ACR 2016 Cr; and 73.8%, 91.7%, 81.8% and 8.94 for AAPT Cr. The best performance in terms of correct classification and LR+ was therefore documented for ACR 2011 Cr.

Fig. 3 PDS *vs* the FAS 2019 modCr



Receiver operating characteristic curve analysis comparing the diagnostic accuracy of PDS *vs* the FAS 2019 modCr. PDS: polysymptomatic distress scale; FAS 2019 modCr: Fibromyalgia Assessment Status 2019 modified Criteria.

Fig. 3 shows the ROC curve analysis for the FAS 2019 modCr and the PDS, which was carried out to assess the ability to discriminate between patients with FM and non-FM for multiple cut-off points. The area under the ROC curve for FAS 2019 modCr was 0.924 (95% CI: 0.903, 0.942), whereas for the PDS it was 0.927 (95% CI: 0.906, 0.945). Instruments showed similar performance (difference between areas =

Table 2 Multiple diagnostic cut-off points of FAS 2019 modCr resulting from ROC curve analysis

Criterion	Sensitivity	95% CI	Specificity	95% CI	LR+	LR−
>6	99.51	98.2, 99.9	14.98	11.3, 19.3	1.17	0.03
>7	99.01	97.5, 99.7	17.43	13.5, 22.0	1.20	0.05
>8	99.01	97.5, 99.7	22.02	17.6, 26.9	1.27	0.04
>9	98.52	96.8, 99.5	28.75	23.9, 34.0	1.38	0.05
>10	98.52	96.8, 99.5	37.92	32.6, 43.4	1.59	0.03
>11	98.02	96.1, 99.1	41.90	36.5, 47.4	1.69	0.04
>12	97.28	95.2, 98.6	51.07	45.5, 56.6	1.99	0.05
>13	96.30	94.0, 97.9	55.35	49.8, 60.8	2.16	0.06
>14	95.56	93.1, 97.3	59.33	53.8, 64.7	2.35	0.07
>15	94.32	91.6, 96.4	64.53	59.1, 69.7	2.66	0.08
>16	92.59	89.6, 94.9	69.11	63.8, 74.1	3.00	0.11
>17	91.11	87.9, 93.7	74.92	69.9, 79.5	3.63	0.12
>18	89.14	85.7, 92.0	81.04	76.4, 85.1	4.70	0.13
>19	86.42	82.7, 89.6	85.32	81.0, 89.0	5.89	0.16
>20*	84.20	80.3, 87.6	88.99	85.1, 92.2	7.65	0.18
>21	82.22	78.1, 85.8	90.83	87.2, 93.7	8.96	0.20
>22	78.02	73.7, 82.0	93.27	90.0, 95.7	11.60	0.24
>23	76.30	71.8, 80.4	94.50	91.4, 96.7	13.86	0.25
>24	73.58	69.0, 77.8	95.11	92.2, 97.2	15.04	0.28
>25	69.88	65.2, 74.3	95.41	92.5, 97.4	15.23	0.32
>26	63.95	59.1, 68.6	95.72	92.9, 97.6	14.94	0.38
>27	59.01	54.0, 63.8	96.33	93.7, 98.1	16.08	0.43
>28	49.14	44.2, 54.1	97.55	95.2, 98.9	20.08	0.52
>29	44.44	39.5, 49.4	97.86	95.6, 99.1	20.76	0.57
>30	40.49	35.7, 45.5	98.17	96.0, 99.3	22.07	0.61
>31	33.83	29.2, 38.7	98.78	96.9, 99.7	27.65	0.67
>33	24.44	20.3, 28.9	98.78	96.9, 99.7	19.98	0.76
>34	18.77	15.1, 22.9	99.39	97.8, 99.9	30.68	0.82

FAS 2019 modCr: Fibromyalgia Assessment Status 2019 modified Criteria; ROC: receiver operating characteristic; CI: confidence interval; LR+: positive likelihood ratio; LR−: negative likelihood ratio; *: optimal cut-off point.

0.0026; $P = 0.759$). The optimal cut-off value that provided a maximal diagnostic accuracy for FAS 2019 modCr, obtained from the ROC analysis, was ≥20 (Youden's index). Based on this cut-off value, the sensitivity was 84.2% (95% CI: 80.3, 87.6), the specificity was 89.0% (95% CI: 85.1, 92.2) and the LR+ was 7.65 (Table 2).

Discussion

In this study we have demonstrated a substantial agreement among the main sets of classification criteria available for FM. However, considering the differences that have emerged, it can be said that the ACR 2011 Cr seem to be the best performing, when compared with clinical judgment, while the AAPT Cr are the worst.

On the other hand, considering the second aim of our study, good properties of the FAS 2019 modCr were found in terms of high sensitivity and specificity, although these were slightly lower when compared with those of PDS.

Over many decades, there have been many efforts to create diagnostic criteria for the condition we now recognize as FM [21, 22]. The multifaceted symptoms and comorbidities associated with FM make it difficult to diagnose, and the disease is still underdiagnosed and undertreated [2, 23]. The diagnosis of FM might take more than 2 years, with patients seeing an average of 3.7 different physicians during that time [24]. Many health care providers, particularly in primary care, report unclear diagnostic criteria, a lack of confidence in using existing criteria for diagnosis, insufficient training or skill in diagnosing FM, and a lack of knowledge of treatment options [25–28].

With the publication of the ACR 2010 Cr and ACR 2011 Cr, the definition of FM moved from a predominantly chronic pain condition to a multi-symptom disorder, and the tender point exam has been eliminated as a requirement for diagnosis [6, 7]. Although the authors of the 2010/2011 criteria re-emphasized the importance of associated symptoms, there may have been too much movement away from chronic pain as the core symptom of FM [26, 27]. Studies of alternative criteria evaluated a variety of associated symptoms along with various definitions of widespread pain in the diagnosis of FM [28, 29]. The authors of the revised ACR 2016 Cr addressed the problem with the 2010/2011 criteria regarding misclassification of patients who did not have generalized pain [30], which occurred because the 2010/2011 criteria did not consider the spatial distribution of painful sites. The ACR

2016 Cr require that patients have pain in four of five regions, called 'generalized pain', to distinguish it from the 1990 definition of 'widespread pain' [9]. Even though there are different definitions of widespread pain and associated symptoms, most of the previous FM criteria appear to identify a similar group of patients that most clinicians would agree have FM.

Arnold and coworkers recently identified, along with chronic pain, fatigue and sleep problems as the two principal associated symptoms, proposing a reduction in the number of non-pain symptoms included as core diagnostic criteria, for several reasons [10]. First, to reduce the complexity of diagnosis and make the FM criteria easier to use in practice. Second, these symptoms occur in the majority of patients with FM. Third, pain, sleep disturbance and fatigue were identified by OMERACT as core symptoms of FM [31].

Moreover, these criteria recommend to clinicians to evaluate the presence of other disorders in order to start the appropriate treatment [10]. This can be challenging in clinical practice, because comorbid disorders, including other chronic pain conditions, are common in patients with FM. A variety of disorders can mimic FM, such as hypothyroidism, and also different medications may contribute to the development of generalized pain, such as statins, aromatase inhibitors, bisphosphonates and opioids (i.e. opioid-induced hyperalgesia). Many other painful diseases (e.g. rheumatoid arthritis, osteoarthritis, spinal stenosis, and mood and anxiety disorders) may co-occur in patients with FM [32].

The clinician must determine the possible contribution of various disorders to the patient's presentation. The presence of other disorders does not necessarily exclude a diagnosis of FM, and all disorders will need clinical attention.

In summary, the problem of identifying the best diagnostic criteria in a heterogeneous syndrome like FM is, at this point, without a specific and definite solution. The development of specific biomarkers or neuroimaging models will probably make the problem simpler. We realized that, from a clinical point of view, we must rely on diffuse chronic muscoloskeketal pain with the inclusion at least of sleep and fatigue as representative of the constellation of the multi-symptom aspects of this syndrome in order not to sink in a myriad of different and confusing symptoms. In this sense, the present study supports the use of the ACR 2011 Cr.

In addition, this study also showed that the FAS 2019 modCr can be useful to evaluate symptom severity according to the clinical phenotype in a similar way to that of the PDS [12]. The usefulness of the FAS 2019 modCr allows a simple and fast evaluation of the diagnosis and of the severity of the disease. We believe the non-articular pain sites should be evaluated by patients during the physician assessment at the time of the first visit, in order to prevent the overdiagnosis of FM [33], but may be used after the diagnosis has been confirmed on a self-evaluation basis of the disease severity by patients themselves [34].

Among the limits of this study we must surely highlight the cross-sectional evaluation, as it is well known that the symptoms of FM can fluctuate widely and for each patient we have only one assessment at one time point, the recruitment from a single centre, and the fact that the diagnostic judgment among clinicians, although experts, can be variable. From a methodological point of view, a limitation is represented by the comparison through the analysis of ROC curves only between FAS 2019 Cr and PDS, without a direct comparison also with ACR 2011 Cr and AAPT Cr.

In conclusion, among the various sets of classification criteria available for FM, the ACR 2011 Cr are those that are most closely in agreement with the clinician's judgement, while the AAPT Cr are the most distant. As regards the FAS 2019 modCr, based on patients' assessment of fatigue, sleep disturbances and pain evaluated on 19 non-articular sites, the operating characteristics are similar to those of more validated instruments (i.e. PDS), with somewhat better ease-of-use, and could be used for diagnosis and follow-up of FM patients.

Acknowledgements

F.S., M.D.C. and S.F. made substantial contributions to the acquisition, analysis and interpretation of data. F.A., D.B., J.N.A., W.H. and P.S.-P. made substantial contributions to the conception and design of the work. F.S., M.D.C. and S.F. drafted the work, and F.A., D.B., J.N.A., W.H. and P.S.-P. revised it critically for important intellectual content. All the authors approved the final version to be published. All the authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Funding: No specific funding was received from any funding bodies in the public, commercial or not-for-profit sectors to carry out the work described in this manuscript.

Disclosure statement: The authors have declared no conflicts of interest.

References

1 Aggarwal R, Ringold S, Khanna D *et al*. Distinctions between diagnostic and classification criteria? Arthritis Care Res (Hoboken) 2015;67:891–7.

2 Choy E, Perrot S, Leon T *et al*. A patient survey of the impact of fibromyalgia and the journey to diagnosis. BMC Health Serv Res 2010;10:102.

3 Salaffi F, De Angelis R, Grassi W; MArche Pain Prevalence; INvestigation Group (MAPPING) study. Prevalence of musculoskeletal conditions in an Italian population sample: results of a regional community-based study. I. The MAPPING study. Clin Exp Rheumatol 2005;23:819–28.

4 Sarzi-Puttini P, Atzeni F, Salaffi F *et al*. Multidisciplinary approach to fibromyalgia: what is the teaching? Best Pract Res Clin Rheumatol 2011;25:311–9.

5 Wolfe F, Smythe HA, Yunus MB *et al*. The American College of Rheumatology 1990 criteria for the classification of fibromyalgia. Report of the multicenter criteria committee. Arthritis Rheum 1990;33:160–72.

6 Wolfe F, Clauw DJ, Fitzcharles MA *et al*. The American College of Rheumatology preliminary diagnostic criteria for fibromyalgia and measurement of symptom severity. Arthritis Care Res (Hoboken) 2010;62:600–10.

7 Wolfe F, Clauw DJ, Fitzcharles MA *et al*. Fibromyalgia criteria and severity scales for clinical and epidemiological studies: a modification of the ACR preliminary diagnostic criteria for fibromyalgia. J Rheumatol 2011;38:1113–22.

8 Wolfe F, Brähler E, Hinz A, Häuser W. Fibromyalgia prevalence, somatic symptom reporting, and the dimensionality of polysymptomatic distress: results from a survey of the general population. Arthritis Care Res (Hoboken) 2013;65:777–85.

9 Wolfe F, Clauw DJ, Fitzcharles MA *et al*. 2016 revisions to the 2010/2011 fibromyalgia diagnostic criteria. Semin Arthritis Rheum 2016;46:319–29.

10 Arnold LM, Bennett RM, Crofford LJ *et al*. AAPT diagnostic criteria for fibromyalgia. J Pain 2019;20:611–28.

11 Okifuji A, Turk DC, Sinclair JD, Starz TW, Marcus DA. A standardized manual tender point survey. I. Development and determination of a threshold point for the identification of positive tender points in fibromyalgia syndrome. J Rheumatol 1997;24:377–83.

12 Salaffi F, Sarzi-Puttini P, Girolimetti R *et al*. Development and validation of the self-administered Fibromyalgia Assessment Status: a disease-specific composite measure for evaluating treatment effect. Arthritis Res Ther 2009;11:R125.

13 Cash E, Boktor SW. Understanding Biostatistics Interpretation. In: StatPearls. StatPearls Publishing, Treasure Island (FL); 2019.

14 Salaffi F, Franchignoni F, Giordano A *et al*. Psychometric characteristics of the Italian version of the revised Fibromyalgia Impact Questionnaire using classical test theory and Rasch analysis. Clin Exp Rheumatol 2013;31(6 Suppl 79):S41–9.

15 Burckhardt CS, Clark SR, Bennett RM. The Fibromyalgia Impact Questionnaire: development and validation. J Rheumatol 1991;18:728–33.

16 Sarzi-Puttini P, Atzeni F, Fiorini T *et al*. Validation of an Italian version of the Fibromyalgia Impact Questionnaire (FIQ-I). Clin Exp Rheumatol 2003;21:459–64.

17 Bennett R. The Fibromyalgia Impact Questionnaire (FIQ): a review of its development, current version, operating characteristics and uses. Clin Exp Rheumatol 2005;23(Suppl 39):S154–62.

18 Salaffi F, Di Carlo M, Arcà S, Galeazzi M. Categorisation of disease severity states in fibromyalgia: a first step to support decision-making in health care policy. Clin Exp Rheumatol 2018;36: 1074–81.

19 Swetz JA. Measuring accuracy of diagnostic systems. Science 1988;240:1285–93.

20 Hanley JA, McNeil BA. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. Radiology 1983;148:839–43.

21 Sarzi-Puttini P, Atzeni F, Masala IF *et al*. Are the ACR 2010 diagnostic criteria for fibromyalgia better than the 1990 criteria? Autoimmun Rev 2018;17:33–5.

22 Salaffi F, Sarzi-Puttini P. Old and new criteria for the classification and diagnosis of fibromyalgia: comparison and evaluation. Clin Exp Rheumatol 2012;30:3–9.

23 Hadker N, Garg S, Chandran AB *et al*. Primary care physicians' perceptions of the challenges and barriers in the timely diagnosis, treatment and management of fibromyalgia. Pain Res Manag 2011;16:440–4.

24 Arnold LM, Stanford SB, Welge JA, Crofford LJ. Development and testing of the fibromyalgia diagnostic screen for primary care. J Womens Health (Larchmt) 2012;21:231–9.

25 Arnold LM, Hudson JI, Keck PE *et al*. Comorbidity of fibromyalgia and psychiatric disorders. J Clin Psychiatry 2006;67:1219–25.

26 Jones GT, Atzeni F, Beasley M *et al*. The prevalence of fibromyalgia in the general population: a comparison of the American College of Rheumatology 1990, 2010, and modified 2010 classification criteria. Arthritis Rheumatol 2015;67:568–75.

27 Dean LE, Arnold L, Crofford L *et al*. Impact of moving from a widespread to multisite pain definition on other fibromyalgia symptoms. Arthritis Care Res (Hoboken) 2017;69:1878–86.

28 Bennett RM, Friend R, Marcus D, Bernstein C *et al*. Criteria for the diagnosis of fibromyalgia: validation of the modified 2010 preliminary American College of Rheumatology criteria and the development of alternative criteria. Arthritis Care Res (Hoboken) 2014;66: 1364–73.

29 Salaffi F, Mozzani F, Draghessi A *et al*. Identifying the symptom and functional domains in patients with fibromyalgia: results of a cross-sectional Internet-based survey in Italy. J Pain Res 2016;9:279–86.

30 Egloff N, von Känel R, Müller V *et al*. Implications of proposed fibromyalgia criteria across other functional pain syndromes. Scand J Rheumatol 2015;44:416–24.

31 Mease P, Arnold LM, Choy EH *et al*. Fibromyalgia syndrome module at OMERACT 9: domain construct. J Rheumatol 2009;36:2318–29.

32 Atzeni F, Cazzola M, Benucci M *et al*. Chronic widespread pain in the spectrum of rheumatological diseases. Best Pract Res Clin Rheumatol 2011;25: 165–71.

33 Häuser W, Sarzi-Puttini P, Fitzcharles MA. Fibromyalgia syndrome: under-, over- and misdiagnosis. Clin Exp Rheumatol 2019;37(1 Suppl 116):90–7.

34 Salaffi F, Ciapetti A, Gasparini S *et al*. Web/Internet-based telemonitoring of a randomized controlled trial evaluating the time-integrated effects of a 24-week multi-component intervention on key health outcomes in patients with fibromyalgia. Clin Exp Rheumatol 2015;33(1 Suppl 88):S93–101.