Running head: FM Diagnostic Criteria

Title: Criteria for the Diagnosis of Fibromyalgia: Validation of the Modified 2010 Preliminary ACR Criteria and the Development of Alternative Criteria

Robert Bennett[1], Ronald Friend[2], Dawn Marcus[3], Cheryl Bernstein[3], Bobby Kwanghoon Han[4], Ralph Yachou[4], Atul Deodar[1], Alan Kaell[5], Peter Bonafede[6], Allan Chino[1] and Kim Jones[1]

Supported by: NIH/NIAMS R21NR012572-01, NIH/NIAMS/1R01AR059102-01A1,NIH/ NIAMS 1R21AR056751-01, NIH/NIAMS 1R21AR053597 - 01A2, NIH UL1 RR024153 and UL1TR000005.

1 Robert M Bennett, MD, Kim Jones, PhD, Atul Deodar, MD, Allan Chino, PhD: Oregon Health & Science University, Portland, OR 97239, USA
2 Ronald Friend, PhD: Department of Psychology: Stony Brook University, Stony Brook, NY 11794, USA
3 Dawn Marcus, MD*, Cheryl Bernstein, MD: Department of Anesthesiology, University of Pittsburgh, Pittsburgh, PA 15261, USA
4 Bobby Han, MD, Ralph Yachou, MD: Cooper Medical School of Rowan University, Camden, NJ 08103, USA
5 Alan Kaell, MD: Rheumatology Associates of Long Island and Stony Brook University, Stony Brook, NY 11794, USA
6 Peter Bonafede, MD: Providence Arthritis Center, Portland, OR 97213

** Deceased*

Body of manuscript: 3,658 words - 3,800 is maximum permitted
Abstract: 247 words - 250 permitted
Figure Legend: 162 words
Number of tables: 5
Number of figures: 1
Number of references: 30

Address correspondence to: Robert Bennett MD, Oregon Health & Science University, 6119 Southwest California Street, Portland, OR 97219. Email: bennetrob1@comcast.net

ABSTRACT

Objective. To validate the 2011 modification of the 2010 American College of Rheumatology (ACR) Preliminary Criteria for the Diagnosis of Fibromyalgia (2011ModCr) and develop alternative criteria, in a sample of patients with diverse pain disorders that are commonly seen in everyday practice by pain specialists, rheumatologists and psychologists.

Methods. Eight clinicians, from geographically varied locations in the United States, evaluated chronic pain and psychiatric patients with a standard set of questions that included the 2011ModCr questions, the Symptom Impact Questionnaire (SIQR), a 28 area pain location inventory (PLI) and the SF-36. Alternative diagnostic criteria were developed from the same data set using logistic regression and receiver operating curve analysis.

Results. Complete data on 321 patients were evaluated; there were 135 with FM (ACR 1990 criteria) and 186 with 16 other common chronic pain problems. Comparing the 2011ModCr with the ACR 1990 criteria provided a sensitivity of 83%, specificity of 67% and a correct classification of 74%. Alternative Criteria were derived from the 10-item symptom score from the SIQR symptoms and the 28 PLI. Maximal diagnostic accuracy was obtained with pain sites ≥ 17 (range 0-28) and SIQR symptom score of ≥ 21 (range 0-50). These alternative criteria (2013 AltCr) had a diagnostic sensitivity of 81%, specificity of 80% and a correct classification of 80%.

Conclusion. The 2011ModCr had robust operating characteristics. Alternative criteria (2013 AltCr), based on symptom items from the SIQR and pain locations from the PLI, had comparable operating characteristics with somewhat better specificity and ease-of-use.

Significance and Innovations:

- Validation of the 2011 Modification of the 2010 ACR Preliminary Diagnostic Criteria for Fibromyalgia by independent group and clinical investigators in a more diverse set of chronic pain conditions than previously reported
- Development of Alternative Criteria for the diagnosis of fibromyalgia using the same set of chronic conditions and clinical investigators
- The alternative criteria have the same sensitivity and somewhat better specificity than the 2011 Modification of the 2010 ACR Preliminary Diagnostic Criteria
- The alternative criteria have improved ease-of-use compared to the 2011 Modification of the 2010 ACR Preliminary Diagnostic Criteria, in that they employ just one combination of pain locations and one time interval for reporting symptoms

The 1990 American College of Rheumatology (ACR) FM Classification Criteria (1) have served a useful purpose in identifying a moderately homogeneous group of chronic pain patients that can

be defined in terms of clinical profiles, aberrant pain physiology and neuroimaging (2). However, it has become increasingly apparent over the last two decades that the 1990 ACR criteria have limited usefulness as regards the diagnosis of FM in clinical practice and cannot be used in epidemiological surveys (3-5). Apart from the fact that the designated 18 location tender point count is seldom used or incorrectly performed (5), it is readily apparent that FM is more than just a pain disorder(6;7).

In 2010 Wolfe and colleagues published a set of clinical diagnostic criteria based on empiric data that included 19 pain locations and 41 somatic symptoms, the "*2010 ACR Preliminary Diagnostic Criteria"* (8). In 2011 these criteria were modified by eliminating the 41 physician-derived estimate of somatic symptoms and substituting 6 self-reported symptoms (impaired sleep, fatigue, poor cognition, headaches, depression and abdominal pain (9). These modified criteria, henceforth referred to as the "2011ModCr" are based on satisfying one of two combinations of Widespread Pain Index (WPI) and Symptom Severity (SS) scores: 3-6 WPI and ≥ 9 SS or ≥ 7 WPI and ≥ 5 SS. A summation of the WPI and SS, the Fibromyalgia Symptom Scale (FS), had a diagnostic cut-off at ≥ 13. These modified criteria classified 60% of FM patients out of a pool of patients with FM, osteoarthritis, systemic lupus erythematosus, fibromyalgia and rheumatoid arthritis. A Japanese version evaluated 462 patients with FM (ACR 1990) and 231 non-FM patients (mainly rheumatoid arthritis), reported a diagnostic sensitivity of 64% and specificity of 96%, using an FS cutoff of ≥ 10 (10). A German version reported a 69 % diagnostic concordance with ACR 1990 criteria (11). Marcus reported on 337 chronic pain patients (mainly arthritis, FM, migraine, neuropathic pain and myofascial pain) who self-screened using the 2011ModCr (12). Using an FS score ≥13, there was a diagnostic sensitivity of 76 % and a specificity of 82 %, when compared with a clinician's diagnosis of FM (12).

These studies report a fairly wide variation in sensitivity, specificity, diagnostic standards, and spectrum of pain disorders that were sampled. In this current study we have evaluated the validity of the 2011ModCr with respect to the ACR 1990 definition in a cross-sectional survey of chronic pain patients seen in the everyday practices of 8 physicians. In addition, an alternative set of diagnostic criteria has been explored.

METHODS AND SUBJECTS

This study was conducted from July 2013 to March 2013. The study protocol was approved by local Institutional Review Boards at participating sites as an exempt study not requiring informed

consent from participants. All participants were informed of the research nature of this project. Neither investigators nor subjects were compensated for participation.

Study subjects

The dominating principle of this study was to incorporate a wide range of common pain disorders as seen in everyday clinical practice (13;14). Subjects were recruited from the practices of 5 rheumatologists, 2 pain specialists and 1 psychologist, with locations in Cooper Medical School Camden NJ, Rheumatology Associates Long Island NY, Oregon Health & Science University Portland OR (OHSU), Providence Medical Center Portland OR, University of Pittsburgh Medical Center and a private psychology practice in Portland OR, respectively. The psychologist only entered subjects with major depressive disorder. The participating clinicians were personally known to the study coordinators (RB and RF) as experienced in the diagnosis and management of FM patients. All subjects were personally evaluated by the site investigator prior to completing an online questionnaire.

Adult patients ≥18 years old were enrolled as a convenience sample with individual investigators inviting patients to participate if they had one or more of the specified diagnoses agreed upon during the development of the study (Table 1). Enrollment was not restricted by sex, comorbidities, medications or disease severity. All physicians used the 1990 ACR classification criteria for the diagnosis of FM; only if the subject was being seen for the first time was another tender point evaluation performed. Major depressive disorder (MDD) was based on DSM-IV. All other diagnoses were based on published guidelines.

Questionnaires

Electronic data collection was conducted through SurveyMonkey**.** The study investigator answered the first two questions, identifying the site and selecting patient diagnosis or diagnoses. All other questions were completed by the subject in the following 5 questionnaires:

(1) Demographics (age, gender, educational level, work status, marital status, number of years with chronic pain and other chronic pain disorders).

(2) The 2011 Modified Criteria for FM (2011 ModCr) (9)

(4) The Short Form 36 (SF-36) (15)

(3) The Symptom Impact Questionnaire (SIQR) (16). *The SIQR is identical to the FIQR with the*

exception that the 3 domains do not use the word "FM", this allows completion by patients with non-FM disorders(16)

(5) A 28 anatomical location inventory rated pain and tenderness on a 0 to 10 severity scale. The 28 locations modified from the Revised FM Impact Questionnaire (FIQR) (18) were: jaw, neck, mid upper back, front of the chest, mid lower back, upper back, lower back, shoulders, arms, hands, wrists, hips, thighs, knees, ankles and feet. The 2 following scores were obtained from these data:

1. Number of pain locations , the pain location inventory (PLI) (range 0-28)
2. Summated intensity of pain at 28 locations (0-280).

The pain question was: "Please estimate your level of pain, over the past 7 days, in the following locations". Anchors were "no pain" and "extremely painful".

We also evaluated the patient's experience of tenderness at the same locations, providing the following 2 scores:

1. Number of tender locations (0-28).
2. Summated intensity of tenderness at 28 locations (0-280).

The tenderness question was: "Please estimate your level of tenderness, over the past 7 days, in the following locations, when touched or pressed". Anchors were "no tenderness" and "extremely tender".

Data analysis

All data were analyzed in *STATISTICA* (version 10) using item analysis and Cronbach alpha. An alternative diagnostic criteria algorithm was explored using the 28 area pain location inventory (PLI) and the 10-item symptom sub-score from the SIQR (16). The PLI and SIQR correlated 0.62, with Cronbach alphas of 0.95 and 0.89, respectively. These 2 continuous predictors were entered simultaneously into a logistic regression (evaluated by the Wald statistic) to produce the receiver operating curve (ROC) to determine the optimal threshold for sensitivity and specificity with respect to ACR 1990 criteria, using the Youden index (sensitivity + specificity - 1) (17). The Hosmer-Lemeshow test was used to evaluate goodness of fit (18), with a non significant finding indicating good model fitness. Chi-square, Fisher exact tests and t-tests were used for evaluating means and proportions, with a primary focus of comparing false positives with true positives on demographic and clinical features, adopting a 0.001 alpha level. All significance tests are two-tailed.

Two separate classification analyses using Random Forest - an iterative bootstrapping classification

tree algorithm that ranks predictor variable importance (19;20) were conducted to rank the importance of specific symptoms and pain locations with respect to ACR 1990 classification: one for the WPI and 6 SS symptoms, the second for the pain location inventory (PLI) and 10 SIQR symptoms.

RESULTS

Demographic findings

The demographics were fairly typical of chronic pain patients; they were middle-aged, predominantly female and their duration of pain symptoms was about 15 years (Table 1). Patients without FM were more likely to be older males. Median educational level was 3, indicating that most subjects had attended some college. The most common pain comorbidities in both groups were osteoarthritis, rheumatoid arthritis, low back pain and upper back pain. Migraine headaches and major depressive disorder were more common in FM patients, whereas peripheral neuropathy was more common in the non-FM group. A non- FM related pain disorder increased the total SIQR score by approximately 10%; however, having a related medical disorder did not significantly affect the total SIQR score. As expected, the number and intensity of pain, tender sites and symptoms reported by patients in the last 7 days were substantially higher in the FM than non-FM group; thus supporting the investigators' diagnosis of FM. Some 51.1% of the FM patients reported a concomitant medical diagnosis; the 5 most common being diabetes (8.1 %), cancer (6.7%), asthma (6.7%), heart disease (5.2%) and hypertension (4.4%). Only 16.2% of FM patients were without at least one other medical disorder or a non- FM pain disorder.

Validation of the 2011 Modified Criteria

A comparison of 2011ModCr and ACR 1990 criteria is shown in Table 2, with true positivity (TP), false positivity (FP), false negativity (FN) and absence of FM (TN) determined with reference to the 1990 ACR classification criteria:

Sensitivity = TP/TP +FN= 112/135 = 83%

Specificity = TN/TN+FP= 125/186 = 67%

Correct classification = (TP+TN)/total # subjects = (112+125)/321= 74%

*Diagnostic odds ratio (OR) = (TP **x** TN)/ (FP **x** FN) = 112 **x** 125/61 **x** 23 = 10.0*

Number needed to diagnose (NND) = 1/(sensitivity–(1- specificity)= 1/(0.83–(1- 0.67) = 2

Using the FS recommended cut- off point of 13 produced similar percentages but a higher sensitivity (92%), lower specificity (63%) and a 75% overall correct classification. With this cut off, 37% of ACR negative patients were classified as being false positives (FP). Using the WPI alone with a cut- off of ≥ 10, yielded a sensitivity of 77% and specificity of 80%, with a correct classification of 79%. Combining the SS and WPI with cut-offs of 4 and 10 (defined by the Youden index), did not improve classification over the WPI used alone with respect to ACR 1990 (sensitivity 76%, specificity 81%; OR 13.3). Thus, in this data set the WPI alone provided the best diagnostic accuracy. The WPI and SS correlated 0.61 , with Cronbach alphas of 0.89 and 0.69, respectively.

Development of Alternative Diagnostic FM Criteria (2013AltCr)

Similar to the *2011ModCr* use of 19 WPI pain locations and 6 SS symptoms, we have developed alternative parallel criteria based on a broader set of symptoms and pain locations; namely the 10 SIQR symptom subscale and the 28 PLI. The PLI and SIQR symptom scores were entered into a logistic regression analysis to generate the ROC curve shown in Figure 1. The optimal threshold levels for sensitivity and specificity were provided by a pain site number of 17 and a SIQR symptom score of 21. Alternative criteria based on this partitioning are henceforward referred to as *"2013AltCr"* and are summarized in Table 3. A comparison of the 2013AltCr and the ACR 1990 criteria is shown in Table 4, with TP, FP, FN and TN determined with reference to the 1990 ACR classification criteria:

Sensitivity = 109/135 = 81%

Specificity = 148/186 = 80%

Correct classification = (109+148) /321= 80%

Diagnostic OR = 109 ***x*** *148/38* ***x*** *26= 16.33*

Number needed to diagnose (NND) = 1/((0.81–(1- 0.8))= 1.6

Cut-off points based on Youden index~~,~~ for each of the pain and tenderness measures considered singly, yielded the following:

# Pain sites (i.e. PLI):	*Sensitivity = 83%,*	*Specificity = 72%;*	*OR = 12.55*
Summated pain:	*Sensitivity = 81%,*	*Specificity = 75%;*	*OR = 13.39*
# Tender sites:	*Sensitivity = 83%,*	*Specificity = 74%;*	*OR = 11.83*
Summated tenderness:	*Sensitivity = 80%,*	*Specificity = 75%;*	*OR = 13.61*

Thus adding the SIQR symptoms to the pain site measure (PLI) raised the specificity of the 2013AltCr from 72% to 80% and yielded a correct classification of 80%; it also reduced the number needed to diagnose (NND) to 1.6.

Item comparisons in Table 4 revealed that FP patients differed from TP patients in only 7 of the 27 comparisons, with only 1 at the 0.001 level (i.e. summated tenderness). Thus the 2013AltCr FPs are similar to TPs, with no significant differences between FPs and TPs in number of pain sites and SIQR symptoms, with the exception of environmental sensitivity. There were no significant differences in any of the SF-36 scales.

Finally, we created a composite PLI plus SIQR Symptom Severity Scale (PLI- SIQR), in a similar vein to the FS, that combines the WPI and the SS. We thus added a patient's PLI score (0- 28) to their SIQR symptom score (0- 50) to produce a PLI- SIQR score (range 0 -78). Logistic regression produced an AUC of 0.83, a sensitivity of 70% and specificity 82% (OR 11.01), with reference to ACR criteria. A PLI- SIQR score of 44 produced a sensitivity of 97%, specificity of 87% (OR 196.32), with respect to the 2013AltCr criteria. The PLI-SIQR provides a classification criterion in a single number (i.e. ≥ 44). The PLI-SIQR has a Cronbach alpha of 0.90. The PLI- SIQR and the FS correlated with SF-36 Pain, 0.68 and 0 .56, respectively.

Clinical features of 2013AltCr positive patients not fulfilling the 1990 ACR criteria

The 2011ModCr embrace a wider spectrum of patients with FM- like symptoms than the ACR 1990 criteria. Those patients diagnosed as having FM with the 2011ModCr who do not fulfill the 1990 ACR criteria were referred to as false positives (FP). Likewise, there is a corresponding group of thirty-eight 2013AltCr FP patients (Table 4). Overall, they had lower scores than TPs on WPI, cognitive problems and FS. On the other hand, they were similar to the TPs on SIQR items, but lower on summated pain, number of tender sites and summated tenderness. There was no significant difference on SF-36 subscales.

Comparison of the 2011ModCr with the 2013AltCr

The 2011ModCr classified FM more frequently than the 2013AltCr (173 versus 147), with 61 (35.3%) and 38 (25.8 %) FPs respectively (Tables 2 and 4). Comparing the FPs from both the 2011ModCr and the 2013AltCr shows a general uniform agreement, except that FPs had significantly lower scores on summated pain and tenderness. Comparing the differences between FPs and TP on the 27 item and scale comparisons in Table 2, revealed more differences in the

2011ModCr, with 11 items at $p \leq 0.001$ level; this compares to just 1 item in the 2013AltCr in Table 4 (Fisher exact test; $p = 0.002$). Thus patients diagnosed as having FM using the 2013AltCr are more similar to ACR 1990-defined patients than 2011ModCr-defined patients. When the 61 2011ModCr FPs were divided into being either 2013AltCr positive or 2013AltCr negative, 34 were classified as 2013AltCr positive and 27 were 2013AltCr negative. Interestingly, there was a difference between these two groups in the distribution of pain locations: peripheral pain locations were significantly more common in the 2013AltCr positives than the 2013AltCr negatives (9.12 vs 4.70; $p < 0.001$).

A notable finding was a 31% prevalence of males in the 2011ModCr FPs, and a 34% prevalence of males in the 2013AltCr FPs; thus the overall male prevalence (i.e.TP + FP) was 16% for the 2011ModCr and 14% for the 2013AltCr, compared with 6% ACR 1990. The 2013AltCr females and males had similar PLI score (24.8 vs. 23.4) and tenderness locations (22.0 vs 20.4), but differed on summated pain (144. 4 vs. 114.9; $p = 0.016$) and summated tenderness (144 vs. 92.5; $p = 0.0008$), with males reporting less pain and tenderness intensity. This result supports the assessment of the intensity of self-reported tenderness and pain as a potentially useful discriminatory variable in fibromyalgia questionnaires.

Random Forest analyses (Figure 1, panels B and D) compared the relative importance of components of the 2011ModCr and the 2013 AltCr. It is seen that both sets of criteria rated the number of pain sites as most important. However, 3 symptoms of the 2013AltCr that stand out are not part of the 2011ModCr (environmental sensitivity, tenderness and stiffness), were rated as 2nd, 3rd and 5th in importance in the 2013AltCr .

The performance of both 2011ModCr and the 2013AltCr was evaluated in their ability to diagnose fibromyalgia in 5 common chronic pain problems (low back pain, chronic migraine, rheumatoid arthritis, osteoarthritis knees, and upper back pain), and separately for female and male patients. Both sets of criteria provided good levels of diagnostic separation (Table 5). While sensitivity was similar for the two criteria, specificity ranged from 50- 75% for the 2011ModCr and 72- 89% for the 2013 AltCr.

DISCUSSION

This study of a cross-section of chronic pain disorders, seen in the geographically diverse, every-day practices of 8 independent physicians, provides robust evidence that the 2011ModCr have satisfactory diagnostic accuracy with respect to the ACR 1990 criteria. Their diagnostic sensitivity was eighty-four percent and the specificity was sixty-seven percent. The rather low specificity is to be expected, as the 2011ModCr was designed to encompass a wider spectrum of patients with FM (8). This wider spectrum of FM diagnoses was also seen with the use of the 2013AltCr. A notable finding, using both criteria, was an increased prevalence of males: sixteen percent using 2011ModCr and fourteen percent using 2013AltCr; this compares with a six percent male prevalence overall, using the ACR 1990 criteria. Thus the 1990 ACR classification criteria results in the under-reporting of male patients, a conclusion proffered by other investigators (21;22). This finding may be related to the higher experience of tenderness in female patients that we found in the current study. Another clinically salient finding was the common occurrence of other pain related comorbidities in FM patients (Table 1). Only thirteen percent of ACR 1990 FM patients lacked another pain disorder. This is probably an underestimated issue in FM studies (23), as most protocols specify that another "significant pain disorder" is a reason for exclusion. This suggests that the interpretation of "significant pain disorder" is either loosely interpreted or the results of many studies do not mimic the FM patients seen in everyday clinical practice. The presence of a non-FM related pain disorder increased the total SIQR score by approximately ten percent; however having a related medical disorder did not significantly affect the total SIQR score.

Using the same data, we have developed Alternative Diagnostic Criteria (2013AltCr) based on a combination of the twenty eight pain location inventory and the ten symptom items from the Symptom Impact Questionnaire (SIQR). These criteria were equally efficient with somewhat better specificity and a smaller number needed to diagnose (NND) than the 2011ModCr. The 2013AltCr were also marginally more efficient in differentiating common chronic pain disorders from FM, as well as male/female patients (Table 5). The difference in specificity of the 2 classification criteria appears to be due to the cutoff values specified in the 2011ModCr, as the area under the curve for the 2 ROCs is almost identical (Figure 1).

The construct of the 2013AltCr differs from the 2011ModCr in the use of more pain locations (twenty-eight versus nineteen) and a larger range of symptoms (ten versus six). The increased number of pain locations was mainly due to the inclusion of peripheral sites; this was deemed appropriate due to the common involvement of hands and feet in FM (24-26); this may have

contributed to the increased specificity of the 2013AltCr vis-à-vis rheumatoid arthritis (see Table 5). While four of the 2013AltCr symptoms (pain, fatigue, poor sleep, and depression) are also included in the 2011ModCr symptom items, an additional three symptoms in the 2013AltCr (environmental sensitivity, tenderness to touch and stiffness) appeared to be especially important in defining FM (Figure 1).

There are several limitations to this study. Diagnostic group sizes were unequal, which is typical of many in most cross-sectional studies. Only nine patients had ankylosing spondylitis and only five had polymyalgia rheumatica; these two conditions are of obvious importance in the differential diagnosis of FM. Well-established FM patients did not have a repeat tender point count at the time of completing the questionnaire; this was justified as the natural history of FM seldom shows remission (27), a fact corroborated by the average pain duration of 14. 6 years, and the higher number of pain and tenderness locations reported by the 1990 ACR FM patients (Table 1). All questionnaire studies suffer from patient recall error; the seven-day and six-month recall interval in some 2011ModCr symptom questions could be problematical (28). In order to more fully establish the diagnostic accuracy of the 2013AltCr, it will be necessary to independently assess its operating characteristics and reproducibility in a larger more diverse population of patients with chronic symptomatology.

A major strength of this study is that two sets of independently developed criteria were validated using a common diagnostic standard on a common sample of subjects, by investigators who were geographically diverse. The range of conditions included in the study was more diverse than any previous studies using the 2011 Modified Criteria or the Original Preliminary ACR 2010 criteria; for instance the current study included patients with migraine, major depression and peripheral neuropathy. Both the 2011ModCr and the 2013AltCr received validation with fourteen separate analyses conducted on sub-samples. The 2013AltCr criteria are easier to use; there is only one time interval for symptom reporting, and are not dependent on two different combinations of pain location ranges and symptom numbers.

While standardized patient questionnaires should prove useful in postal and online epidemiological studies, their role in the clinic requires further scrutiny. It is generally agreed that diagnosis should not be based solely on statistically determined cutoff values derived from a list of questions. A complete clinical evaluation is required to rule out FM mimics and unearth other issues, such as depression, arthritis and symptomatic myofascial trigger points, which would

modify the approach to treatment. Questionnaires could prove helpful in suggesting a diagnosis of FM, but we advocate their use in a Bayesian context with an a priori estimate based on the face-to-face clinical evaluation (29;30). For instance with a pretest probability of FM at 40%, the 2011ModCr provide a posttest probability of 63%, and the 2013AltCr provide a posttest probability of 73%.

In conclusion, we have provided independent validation of the 2011ModCr in a patient sample composed of a wider range of chronic pain disorders than reported in previous studies. We have also explored alternative criteria (2013AltCr) based on the SIQR symptom scale and number of pain locations. These alternative criteria are comparable to the 2011ModCr in diagnostic sensitivity and somewhat better in specificity; importantly they have the advantage of using just one combination of pain locations and one time interval for reporting symptoms.

DEDICATION

We dedicate this paper to the memory of Dr. Dawn Marcus, in recognition of a professional life of continuing scholarship and the compassionate care of patients with chronic pain.

ACKNOWLEDGMENTS

We would like to acknowledge the clinical research assistants at each of the investigator sites. We thank our patients for their participation in this study.

AUTHOR CONTRIBUTIONS

All authors were involved in drafting the article or revising it critically for important intellectual content. All authors approved the final version to be published. Drs. Bennett and Friend had full access to all the data and take responsibility for the integrity of the data and accuracy of the analysis.

Study conception and design: Bennett, Jones, Friend

Evaluation of subjects: Marcus, Bernstein, Han, Yachou, Deodar, Kaell, Bonafede, Chino

Acquisition of data: Bennett

Analysis and interpretation of data: Friend, Bennett

References

(1) Wolfe F, Smythe HA, Yunus MB, Bennett RM, Bombardier C, Goldenberg DL et al. The American College of Rheumatology 1990 Criteria for the Classification of Fibromyalgia. Report of the Multicenter Criteria Committee. Arthritis Rheum 1990; 33(2):160-72.

(2) Clauw DJ, Arnold LM, McCarberg BH. The science of fibromyalgia. Mayo Clin Proc 2011; 86(9):907-11.

(3) Goldenberg DL. Diagnosis and differential diagnosis of fibromyalgia. Am J Med 2009; 122(12 Suppl):S14-S21.

(4) Choy EH, Arnold LM, Clauw DJ, Crofford LJ, Glass JM, Simon LS et al. Content and criterion validity of the preliminary core dataset for clinical trials in fibromyalgia syndrome. J Rheumatol 2009; 36(10):2330-4.

(5) Katz RS, Wolfe F, Michaud K. Fibromyalgia diagnosis: a comparison of clinical, survey, and American College of Rheumatology criteria. Arthritis Rheum 2006; 54(1):169-76.

(6) Wolfe F, Hauser W. Fibromyalgia diagnosis and diagnostic criteria. Ann Med 2011; 43(7):495-502

(7) Arnold LM, Clauw DJ, McCarberg BH. Improving the recognition and diagnosis of fibromyalgia. Mayo Clin Proc 2011; 86(5):457-64.

(8) Wolfe F, Clauw DJ, Fitzcharles MA, Goldenberg DL, Katz RS, Mease P et al. The American College of Rheumatology preliminary diagnostic criteria for fibromyalgia and measurement of symptom severity. Arthritis Care Res 2010; 62(5):600-10.

(9) Wolfe F, Clauw DJ, Fitzcharles MA, Goldenberg DL, Hauser W, Katz RS 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;1113-22.

(10) Usui C, Hatta K, Aratani S, Yagishita N, Nishioka K, Kanazawa T et al. The Japanese version of the 2010 American College of Rheumatology Preliminary Diagnostic Criteria for Fibromyalgia and the Fibromyalgia Symptom Scale: reliability and validity. Mod Rheumatol 2012; 22(1):40-4.

(11) Hauser W, Jung E, Erbsloh-Moller B, Gesmann M, Kuhn-Becker H, Petermann F et al. Validation of the Fibromyalgia Survey Questionnaire within a cross-sectional survey. PLoS ONE 2012; 7(5):e37504.

(12) Marcus DA, Bernstein C, Albrecht KL. Brief, self-report fibromyalgia screener evaluated in a sample of chronic pain patients. Pain Med 2013; 14(5):730-5.

(13) Whiting P, Rutjes AW, Reitsma JB, Glas AS, Bossuyt PM, Kleijnen J. Sources of variation and bias in studies of diagnostic accuracy: a systematic review. Ann Intern Med 2004; 140(3):189-202.

(14) Leeflang MM, Bossuyt PM, Irwig L. Diagnostic test accuracy may vary with prevalence: implications for evidence-based diagnosis. J Clin Epidemiol 2009; 62(1):5-12.

(15) Jenkinson C, Coulter A, Wright L. Short form 36 (SF36) health survey questionnaire: normative data for adults of working age. BMJ 1993; 306(6890):1437-40.

(16) Bennett RM, Friend R, Jones KD, Ward R, Han BK, Ross RL. The Revised Fibromyalgia Impact Questionnaire (FIQR): validation and psychometric properties. Arthritis Res Ther 2009; 11(4):R120.

(17) Fluss R, Faraggi D, Reiser B. Estimation of the Youden Index and its associated cutoff point. Biom J 2005; 47(4):458-72.

(18) Paul P, Pennell ML, Lemeshow S. Standardizing the power of the Hosmer-Lemeshow goodness of fit test in large data sets. Stat Med 2013; 32(1):67-80.

(19) Strobl C, Malley J, Tutz G. An introduction to recursive partitioning: rationale, application, and characteristics of classification and regression trees, bagging, and random forests. Psychol Methods 2009; 14(4):323-48.

(20) Breiman L. Random Forests. Machine Learning 2001; 45(1):5-52.

(21) Katz JD, Mamyrova G, Guzhva O, Furmark L. Gender bias in diagnosing fibromyalgia. Gend Med 2010; 7(1):19-27.

(22) Hauser W, Kuhn-Becker H, von WH, Settan M, Brahler E, Petzke F. Demographic and clinical features of patients with fibromyalgia syndrome of different settings: a gender comparison. Gend Med 2011; 8(2):116-25.

(23) Yunus MB. The prevalence of fibromyalgia in other chronic pain conditions. Pain Research Treatment 2012;1-8.

(24) Reilly PA, Littlejohn GO. Peripheral arthralgic presentation of fibrositis/fibromyalgia syndrome. J Rheumatol 1992; 19(2):281-3.

(25) Friend R, Bennett RM. Distinguishing fibromyalgia from rheumatoid arthritis and systemic lupus in clinical questionnaires: an analysis of the revised Fibromyalgia Impact Questionnaire (FIQR) and its variant, the Symptom Impact Questionnaire (SIQR), along with pain locations. Arthritis Res Ther 2011; 13(2):R58.

(26) Hauser W, Hayo S, Biewer W, Gesmann M, Kuhn-Becker H, Petzke F et al. Diagnosis of fibromyalgia syndrome-a comparison of Association of the Medical Scientific Societies in Germany, survey, and American College of Rheumatology criteria. Clin J Pain 2010; 26(6):505-11.

(27) Kennedy M, Felson DT. A prospective long-term study of fibromyalgia syndrome. Arthritis Rheum 1996; 39:682-5.

(28) Clarke PM, Fiebig DG, Gerdtham UG. Optimal recall length in survey design. J Health Econ 2008; 27(5):1275-84.

(29) Medow MA, Lucey CR. A qualitative approach to Bayes' theorem. Evid Based Med 2011; 16(6):163-7.

(30) Sox HC, Jr. Probability theory in the use of diagnostic tests. An introduction to critical study of the literature. Ann Intern Med 1986; 104(1):60-6.

Table 1: Sample demographics, specified chronic pain disorders and major study variables

Demographic and clinical features	Total # (N = 321)	Non- fibromyalgia (N = 186)	Fibromyalgia (N = 135)
Sample demographics:			
Age (mean years ± SD)	N.A.	54.38 ± 13.75	48.70 ± 11.80
Marital status (percent married)	N.A.	51.08%	42.54%
Currently employed	N.A.	29.03%	32.59%
Education level (mean ± SD))*	N.A.	3.02 ± 1.17	3.10 ± 0. 99
Females	242	61.83%	94.07%
Males	79	38.17%	5.93%
Specified diagnoses:			
1. Osteoarthritis of knees	51	18.28%	12.59%
2. Osteoarthritis and hips	27	11.83%	3.70%
3. Osteoarthritis of hands	10	3.76%	2.22%
4. Rheumatoid arthritis	55	19.35%	14.07%
5. Systemic lupus erythematosus (SLE)	21	3.76%	10.37%
6. Sjogren's syndrome	13	3.23%	5.19%
7. Ankylosing spondylitis	8	4.30%	0%
8. Psoriatic arthritis	12	5.91%	0.74%
9. Focal myofascial pain	11	5.38%	0.74%
10. Widespread pain (not FM)	21	8.06%	4.44%
11. Chronic low back pain	69	25.27%	16.30%
12. Chronic upper back pain	37	11.83%	11.11%
13. Polymyalgia rheumatica	5	2.69%	0%
14. Chronic migraine headaches	64	12.37%	30.37%
15. Peripheral neuropathy	33	15.05%	3.70%
16. Major depressive disorder	50	8.06%	25.93%
Major study variables:			
Duration of pain (mean years ± SD)	N.A.	14.16 ± 12.63	14.56 ± 11.11
Number of pain sites (0 - 28)	N.A.	12.81 ± 7.93	22.52 ± 6.04
Summated pain score (0 - 280)	N.A.	59.50 ± 47.11	129.59 ± 58.33
Number of tender sites (0 - 28)	N.A.	9.91 ± 7.98	20.82 ± 7.26
Summated tenderness score (0 - 280)	N.A.	53.94 ± 55.67	131.75 ± 67.30
Total SIQR score (0 - 100)	N.A.	41.48 ± 23.33	62.08 ± 20.72
SIQR symptom score (0 - 50)	N.A.	21.35 ± 10.98	32.00 ± 19.10
SS score (0- 12)	N.A.	4.78 ± 2.36	7.29 ± 2.01
WPI score (0 - 19)	N.A.	6.20 ± 4.47	12.42 ± 4.17
FS score (WPI + SS) (0 - 31)	N.A.	10.98 ± 6.02	19.71 ± 5.26

FM= fibromyalgia

All values are means ± 1 S.D.

Note: The fibromyalgia diagnosis was based on the 1990 ACR classification criteria

SIQR scores are equivalent to the FIQR scores in fibromyalgia patients (20)

*Educational level ranged from "some high school" (0) to "graduate degree" (5)

Table 2: Comparison of Wolfe 2011 diagnostic criteria with 1990 ACR classification criteria

	TP 1990ACR(+) 2011ModCr(+)	FP 1990ACR(–) 2011ModCr(+)	FN 1990ACR(+) 2011ModCr(-)	TN 1990ACR(-) 2011ModCr(-)	TP vs. FP p values
Numbers of subjects in each category	**112**	**61**	**23**	**125**	
Demographics					
Age (mean years)	47.26	50.54	55.70	56.26	0.065
Pain duration	14.46	14.62	15.04	13.91	0.772
Gender (percent female)	92.86	68.85	100.00	58.40	<0.0001
Education (mean)	3.07	2.77	3.26	3.14	0.063
Number of pain comorbidities	2.46	1.90	2.17	1.53	0.006
Employed (percent)	32.14	29.51	34.78	28.80	0.721
MDD diagnosis (percent)	29.46	18.03	8.70	3.20	0.105
Wolfe 2011 analyses (mean values)					
WPI (0- 19)	13.25	10.59	8.39	4.06	<0.0001
SS (0- 12)	7.88	6.93	4.39	3.74	<0.0001
Fatigue (0- 3)	2.54	2.34	1.57	1.27	0.065
Un- refreshing sleep (0- 3)	1.93	1.59	0 .61	0.66	0.019
Cognitive problems (0- 3)	0 .98	0.70	0.65	0.40	0.001
Headaches (0- 1)	0.71	0.59	0.35	0.24	0.127
Abdominal pain or cramps(0- 1)	0 .80	0.74	0.30	0.31	0.341
Depression (0- 1)	1.00	0.97	0.91	0.85	0.123
FS (WPI + SS) (0- 31)	21.13	17.52	12.78	7.79	<0.0001
SIQR analyses (mean values)					
Total SIQR score (0- 100)	66.65	60.64	39.83	32.13	0.035
Function sub- score (0- 30)	18.46	17.24	13.00	8.82	0.290
Overall sub- score (0- 20)	13.84	13.26	6.26	6.25	0.503
Symptoms sub- score (0- 50)	34.35	30.14	20.57	17.06	0.001
SIQR Pain (0- 10)	7.46	6.74	6.35	5.07	0.022
SIQR Energy (0- 10)	7.26	6.74	4.96	4.65	0.147
SIQR Stiffness (0- 10)	7.60	7.28	5.30	4.30	0.350
SIQR Sleep (0- 10)	8.45	7.80	6.61	4.96	0.083
SIQR Depression (0- 10)	5.59	5.13	1.78	2.25	0.374
SIQR Memory (0- 10)	6.12	4.89	2.17	2.31	0.012
SIQR Anxiety (0- 10)	6.01	5.70	2.39	2.48	0.547
SIQR Balance (0- 10)	5.60	4.28	3.26	2.26	0.005
SIQR Tenderness (0- 10)	7.59	6.33	5.09	3.46	0.001
SIQR Environmental sensitivity (0- 10)	7.04	5.39	3.22	2.38	0.001
28 Pain Location Point analysis (mean					
Number of pain sites (0- 28)	23.70	20.64	16.78	8.98	<0.0001
Summated pain (0- 280)	139.54	98.30	81.13	40.57	<0.0001
Number of tender sites (0- 28)	21.94	16.20	15.39	6.85	<0.0001
Summated tenderness (0- 280)	141.49	84.77	84.33	38.89	<0.0001

TP = true positive, FP = false positive, FN = false negative, TN = true negative

1990ACR(+) = patients fulfilling the ACR 1990 criteria.

1990 ACR(–) = patients not fulfilling the ACR 1990 criteria

2011ModCr(+) = patients fulfilling the 2011 Modified Criteria

2011ModCr(-)= patients not fulfilling the 2011 Modified Criteria

Table 3: Alternative criteria (2013AltCr)

Pain location inventory (PLI)

Directions: For each of the following 28 sites, select those locations where you have experienced persistent pain during the past 7 days. The score will be between 0 and 28.

☐	☐	☐	☐
☐	☐	☐	☐
☐	☐	☐	☐
☐	☐	☐	☐
☐	☐	☐	☐
☐	☐	☐	☐
☐	☐	☐	☐

10-item SIQR symptoms:

Directions: For each of the following 10 questions, check the one box that best indicates the intensity of the following common symptoms over the last 7 days.

Criteria:

			0	1	2	3	4	5	6	7	8	9	10	
1.	*Pain*	No pain	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	Unbearable pain
2.	*Energy*	Lots of energy	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	No energy
3.	*Stiffness*	No stiffness	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	Severe stiffness
4.	*Sleep*	Awoke rested	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	Awoke very tired
5.	*Depression*	No depression	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	Very depressed
6.	*Memory Problems*	Good memory	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	Very poor memory
7.	*Anxiety*	Not anxious	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	Very anxious
8.	*Tenderness To Touch*	No tenderness	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	Very tender
9.	*Balance Problems*	No imbalance	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	Severe imbalance
10.	*Sensitivity to loud noises, bright lights, odors and cold*	No sensitivity	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	Extreme sensitivity

Note: Summate the 10 individuals scores; the range will be between 0 and 100. Divide this summated score by 2 to obtain the SIQR symptom score.

A patient fulfilling the following guidelines has a high likelihood of having FM:*

1. **The symptoms and pain locations have been persistent for at least the last 3 months**
2. **Pain location score is ≥ 17**
3. **SIQR symptom score is ≥ 21**

** 1. Fibromyalgia patients have a continuum of symptoms; a diagnosis based on a strict numerical cutoff is subject to error.*

** 2. The presence of another pain disorders or related symptoms does not rule out a diagnosis of fibromyalgia.*

** 3. A careful clinical evaluation is always required in order to identify any condition that could fully account for the patient's symptoms and/or contribute to the severity of the symptoms.*

Table 4: Comparison of the 2013AltCrdiagnostic criteria with the 1990 ACR classification criteria

	TP	**FP**	**FN**	**TN**	
	1990ACR(+) 2013AltCr(+)	**1990ACR(–) 2013AltCr(+)**	**1990ACR(+) 2013AltCr(–)**	**1990ACR(–) 2013AltCr(–)**	**TP vs. FP p values**
Numbers of subjects in each category	**109**	**38**	**26**	**148**	
Demographics)					
Age(years)	48.13	49.24	51.08	55.70	0.585
Pain duration	14.77	15.42	13.65	13.84	0.765
Gender (percent female)	93.58	65.79	96.15	60.81	<0.0001
Education (0- 5)	3.07	2.82	3.23	3.07	0.171
Number of pain comorbidities	2.43	1.84	2.35	1.60	0.012
Employed (percent)	28.44	23.68	50.00	30.41	0.571
MDD diagnosis (percent)	29.36	15.79	11.54	6.08	0.132
Wolfe 2011 analyses (mean values)					
WPI (0- 19)	13.39	12.00	8.35	4.71	0.036
SS (0- 12)	7.64	6.92	5.81	4.24	0.033
Fatigue (0- 3)	2.49	2.37	1.88	1.43	0.356
Un- refreshing sleep (0- 3)	1.83	1.55	1.15	0.82	0.108
Cognitive problems (0- 3)	0.87	0.68	0.81	0.45	0.014
Headaches (0- 1)	0.67	0.58	0.58	0.30	0.330
Abdominal pain or cramps(0- 1)	0.79	0.76	0.42	0.37	0.450
Depression (0- 1)	0.99	0.97	0.96	0.86	0.452
FS (WPI + SS) (0- 31)	21.04	18.92	14.15	8.95	0.010
SIQR analyses (mean values)					
Total SIQR score (0- 100)	67.22	63.73	40.52	35.77	0.271
Function sub- score (0- 30)	18.91	17.81	11.72	9.98	0.402
Overall sub- score (0- 20)	13.89	13.47	6.92	7.28	0.670
Symptoms sub- score (0- 50)	34.42	32.45	21.88	18.50	0.156
SIQR Pain (0- 10)	7.55	7.24	6.08	5.20	0.369
SIQR Energy (0- 10)	7.29	7.32	5.08	4.82	0.955
SIQR Stiffness (0- 10)	7.63	7.18	5.42	4.79	0.244
SIQR Sleep (0- 10)	8.44	8.18	6.85	5.30	0.529
SIQR Depression (0- 10)	5.59	5.50	2.23	2.60	0.886
SIQR Memory (0- 10)	5.92	5.53	3.50	2.55	0.504
SIQR Anxiety (0- 10)	6.07	6.03	2.54	2.90	0.937
SIQR Balance (0- 10)	5.59	5.34	3.58	2.30	0.653
SIQR Tenderness (0- 10)	7.72	7.00	4.85	3.73	0.102
SIQR Environmental sensitivity (0- 10)	7.04	5.58	3.65	2.80	0.008
28 Pain Location Point analysis (mean					
Number of pain sites (0- 28)	24.77	24.58	13.08	9.78	0.775
Summated pain (0- 280)	146..39	123.02	59.12	43.19	0.015
Number of tender sites (0- 28)	22.84	19.47	12.35	7.46	0.004
Summated tenderness (0- 280)	147.61	107.18	65.27	40.27	0.001

TP = true positive, FP = false positive, FN = false negative, TN = true negative
1990ACR(+) = patients fulfilling the ACR 1990 criteria.
1990ACR(–) = patients not fulfilling the ACR 1990 criteria
2013AltCr (+) = patients fulfilling the 2011 Modified Criteria
2013AltCr(–) = patients not fulfilling the 2011 Modified Criteria

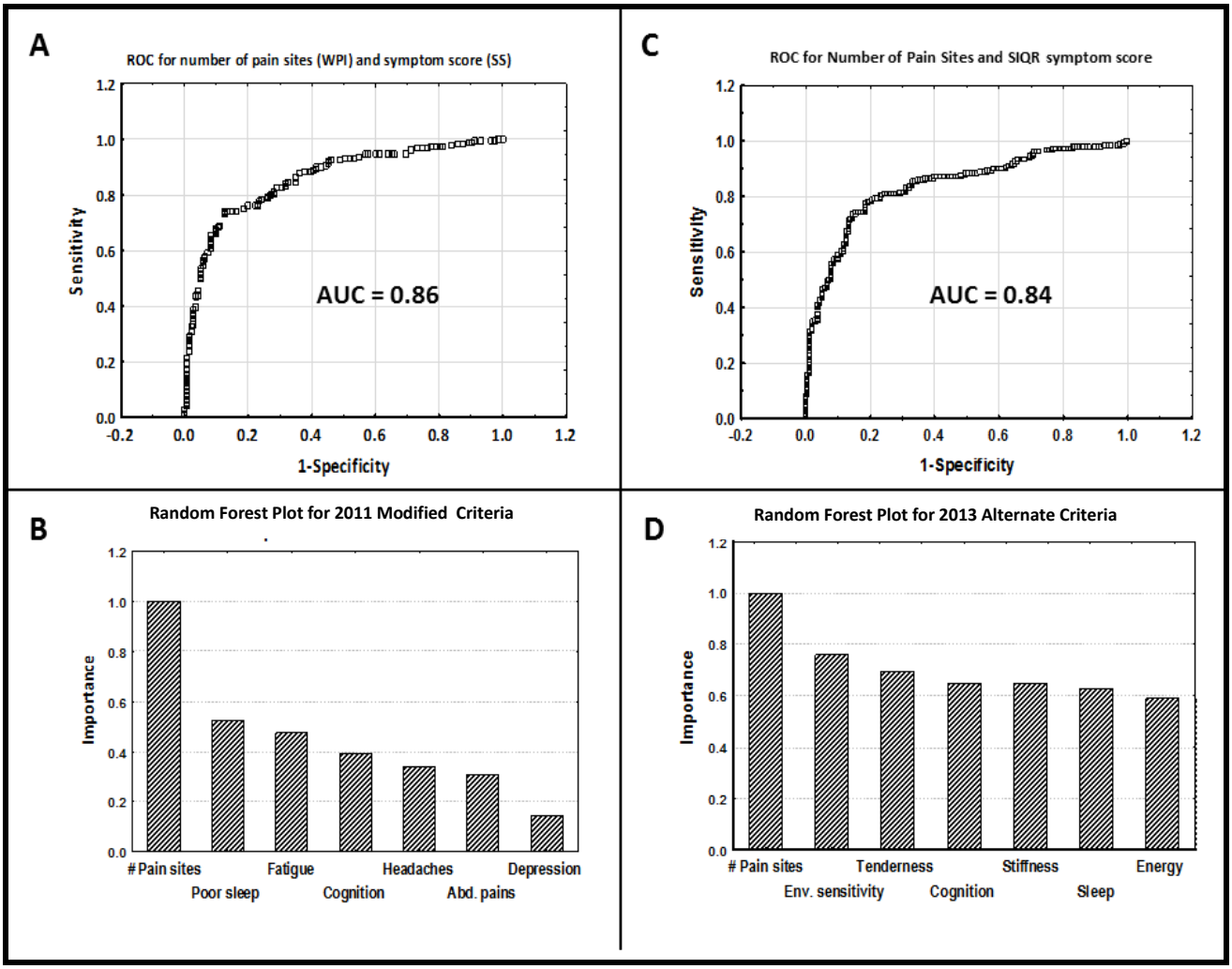
Table 5: Comparison of the 2013AltCr with the 2011ModCr in the diagnosis of fibromyalgia in the total sample, female and male patients , and five different pain disorders

Diagnosis (N)	Diagnostic criteria	Sensitivity	Specificity	% Correct
All patients (321)	*2013AltCr*	80.7	79.6	80.1
	2011ModCr	83.5	67.2	73.8
Female patients (242)	*2013AltCr*	80.3	78.3	79.3
	2011ModCr	81.9	63.5	73.1
Male patients (79)	*2013AltCr*	87.5	81.7	82.3
	2011ModCr	100	73.2	75..9
Low back pain (69)	*2013AltCr*	77.3	72.3	73.9
	2011ModCr	86.4	59.6	68.1
Chronic migraine (64)	*2013AltCr*	78.1	73.9	76.6
	2011ModCr	90.2	65.2	79.7
Rheumatoid arthritis (55)	*2013AltCr*	84.2	88.9	87.3
	2011ModCr	84.2	75.0	78.2
Osteoarthritis knees (51)	*2013AltCr*	64.7	82.4	76.5
	2011ModCr	64.7	73.5	70.6
Upper back pain (37)	*2013AltCr*	86.7	72.7	78.4
	2011ModCr	93.3	50.0	67.6

Sensitivity and specificity for each of the specific pain disorders were calculated by applying the 2011ModCr and 2013AltCr , respectively, to each subgroup with respect to ACR 1990 criteria.

For example, applying the 2013AltCr cut-off points to the 55 RA subjects (2013AltCr + ve, 16; 2013AltCr- ve, 32; ACR + ve, 19, ACR - ve, 36), results in 16/19 or 84.2% sensitivity and 32/36 or 88.9% specificity.

Figure 1



Plots A and C are the receiver operating curves (ROC) for the 2011ModCr and the 2013AltCr in reference to ACR 1990 classification. Plot A provides the ROC curve predicted by the 19 item Widespread Pain Index (WPI) and 6 symptoms SS components of the 2011ModCr. Plot C provides the ROC curve predicted from the 28 Pain Location Inventory (PLI) and the 10 symptoms from the Symptom Impact Questionnaire (SIQR). Both plots are generated by logistic regression. AUC = Area under the curve.

Related statistics: ***A.*** *WPI: Wald=70.37, p=42.41; p<0.0001. SS (Wald Statistic, 14.70, p p<0.0001). Hosmer- Lemeshow Test=4.91, p=0.77. Optimal sensitivity 81%, specificity 77% (*based on WPI=>9 and SS=>4)

C. *PLI: Wald= 40.01; p<0.0001. SIQR symptoms (Wald Statistic, 10.05, p =0. 0015). Hosmer- Lemeshow Test = 8.26, p=0.41. Optimal sensitivity 81%, specificity 80%*

Plots B and D are the Random Forest analyses for the 2011ModCr and the 2013AltCr in predicting ACR classification. They provide information on the relative importance of the individual symptoms and number of pain sites composing the 2011ModCr and the 2013AltCr, respectively.