Comparing three fibromyalgia screening tools: diagnostic accuracy and clinical utility

F. Salaffi[1], M. Carotti[2], S. Farah[1], P. Sarzi-Puttini[3], M. Di Carlo[1]

[1]Rheumatology Unit, Università Politecnica delle Marche, Carlo Urbani Hospital, Jesi (Ancona); [2]Radiology Clinic, Department of Radiology, Università Politecnica delle Marche, Ospedali Riuniti di Ancona; [3]Department of Biomedical and Clinical Sciences, University of Milan - IRCCS Ospedale Galeazzi Sant'Ambrogio, Milan, Italy.

Abstract

Objective

Brief self-administered screening tools have been developed to facilitate rapid recognition of fibromyalgia (FM) features. This study evaluated the diagnostic performance of three commonly used screening tools.

Methods

In this cross-sectional study, consecutive female patients with chronic multisite pain completed three FM screening tools, respectively the Fibromyalgia Rapid Screening Tool (FiRST), the Nociplastic-Based Fibromyalgia Features (NFF) tool, and the Self-Administered Fibromyalgia Screening (SIFIS) questionnaire. Patients were dichotomised in FM (2016 ACR criteria) and non-FM controls, including patients with common non-inflammatory musculoskeletal pain conditions. Discriminative ability was evaluated using the Area Under the Receiver Operating Characteristic Curve (AUC-ROC) analysis and Decision Curve Analysis (DCA) quantified net clinical benefit across probability thresholds.

Results

Among 205 patients, 106 (51.7%) met criteria for FM. SIFIS demonstrated the highest discriminative performance (AUC=0.941) with sensitivity 82.1%, specificity 91.8%, and accuracy 86.8% at a cut-off ≥4. FiRST (cut-off ≥5) achieved AUC=0.827, sensitivity 75.5%, specificity 69.4%, accuracy 77.0%. NFF (cut-off ≥4) achieved AUC=0.829, sensitivity 80.2%, specificity 69.4%, accuracy 75.0%. Differences between SIFIS and the other tools were statistically significant ($p<0.001$). DCA showed that SIFIS provided the highest net clinical benefit across a broad range of thresholds.

Conclusions

In patients with chronic multisite musculoskeletal pain, SIFIS exhibits superior diagnostic accuracy and clinical utility compared with FiRST and NFF. Integration of ROC and DCA demonstrates that SIFIS not only discriminates FM effectively but also offers meaningful clinical benefit, supporting its role as a practical screening tool in routine clinical practice.

Key words

fibromyalgia, screening tools, receiver operating characteristic analysis, decision curve analysis, diagnostic accuracy

Fausto Salaffi, MD, PhD
Marina Carotti, MD
Sonia Farah, Eng
Piercarlo Sarzi-Puttini, MD
Marco Di Carlo, MD

Please address correspondence to:
Prof. Marco Di Carlo
Rheumatology Unit,
Università Politecnica delle Marche,
Carlo Urbani Hospital,
Via Aldo Moro 25,
60035 Jesi (Ancona), Italy.
E-mail address: dica.marco@yahoo.it

Received on April 8, 2026; accepted in revised form on May 25, 2026.

ORCID iDs:
F. Salaffi: 0000-0002-3794-6831
M. Carotti: 0000-0001-6562-180X
S. Farah: 0000-0002-9815-2621
P. Sarzi-Puttini: 0000-0002-8673-5133
M. Di Carlo: 0000-0002-0906-4647

Competing interests: F. Salaffi, S. Farah, M. Carotti and M. Di Carlo are among the authors who developed and validated the SImple FIbromyalgia Screening questionnaire (SIFIS) (ref 27 of the present manuscript).

Introduction

Fibromyalgia (FM) is a common chronic pain disorder characterised by widespread musculoskeletal pain accompanied by fatigue, sleep disturbances and cognitive impairment (1-4). It is currently regarded as a prototypical nociplastic pain condition, reflecting altered central pain processing rather than ongoing peripheral tissue damage (5). Recent epidemiological studies using updated American College of Rheumatology (ACR) 2016 criteria (6) indicate that FM prevalence in European populations remains consistent with earlier estimates, with meta-analytic pooled prevalence ranging from approximately 2-4% (7-9) and 2.22% in the Italian population (10). These aspects position FM among the most prevalent chronic pain conditions in industrialised countries.

Despite the availability of standardised classification criteria, early identification of FM remains challenging. FM continues to be one of the most underdiagnosed and misunderstood disorders in contemporary clinical practice (11, 12). Substantial symptoms overlap with other chronic musculoskeletal pain conditions, including osteoarthritis (OA), mechanical spine syndromes and periarticular disorders, significantly increases the risk of misclassification or diagnostic delay, particularly in non-specialist settings (13, 14). This difficulty is further compounded by the lack of specific biomarkers, longstanding debates surrounding FM's nosological status, and marked interindividual variability in symptom expression, all contributing to persistent diagnostic latency.

The formal diagnostic criteria for FM have undergone major revisions, from the ACR 1990 criteria requiring standardised clinical examination (15), to the ACR 2010 preliminary criteria and the ACR 2016 revision, which operationalised symptom-based assessment through patient-reported measures (6, 16). These developments have facilitated the use of brief, self-administered screening tools capable of identifying individuals at high probability of FM without requiring specialist evaluation. Early diagnosis is not merely a clinical convenience; it is a critical determinant of effective management and improved quality of life (2). On average, patients with FM wait between 2 and 5 years, and often longer, before receiving a definitive diagnosis (3, 12). During this period, many undergo a cycle of specialist consultations, inconclusive investigations and misdiagnoses ranging from inflammatory disease to mood disorders or even somatisation. This prolonged diagnostic uncertainty imposes a substantial psychological burden and may exacerbate symptom severity (12).

Timely diagnosis enables earlier initiation of targeted, multimodal treatment, including patient education, pharmacotherapy, structured physical exercise, cognitive-behavioural therapy, and lifestyle modification (17-19). When implemented early, before symptoms become deeply entrenched, these interventions are associated with substantially improved outcomes (2, 20). Early recognition also allows patients to better understand and contextualise their symptoms. Many individuals report significant relief upon receiving a formal diagnosis (21), and this sense of validation has recognised therapeutic value: it reduces anxiety, improves adherence, and fosters active engagement in disease management (22, 23). Conversely, delayed diagnosis increases the risk of unnecessary investigations, invasive procedures, or inappropriate pharmacological treatment, thereby contributing to patient harm and increased healthcare costs (24).

Primary care physicians are typically the first healthcare professionals consulted by patients with FM and therefore play a pivotal role in early recognition, diagnosis and initial management (2). Screening tools and structured questionnaires can help clinicians identify potential FM cases more consistently.

In the past decade, several concise self-administered screening instruments have been developed to support timely identification of individuals likely to meet FM criteria in routine care. Among these, the Fibromyalgia Rapid Screening Tool (FiRST) (25), the Nociplastic-Based Fibromyalgia Fea-

tures (NFF) tool (26), and the SImple FIbromyalgia Screening questionnaire (SIFIS) (27) are widely used in high-volume outpatient settings. However, comparative evaluations of their diagnostic performance in real-world clinical populations remain limited.

The objective of this study was not only to compare the discriminatory performance of these tools but also to provide a comprehensive assessment of model behaviour using a Bayesian decision-theoretic framework. Receiver Operating Characteristic (ROC) curve analysis (28, 29) was first employed to quantify the discrimination of each instrument independently of decision thresholds. Because discrimination alone does not determine clinical usefulness, Decision Curve Analysis (DCA) was then applied to evaluate whether superior ROC performance (30) translated into greater expected clinical utility when model-predicted probabilities were mapped onto clinically meaningful threshold probabilities. By integrating threshold-independent discrimination (ROC) with threshold-dependent net benefit estimation (DCA), this approach links predictive accuracy with real-world decision consequences.

Methods

Study design and population

This observational cross-sectional study enrolled consecutive female patients presenting for their first visit with suspected FM at the Rheumatology Unit, Università Politecnica delle Marche, Carlo Urbani Hospital, Jesi (Ancona), Italy. Recruitment occurred between October 2024 and March 2026.

Eligible patients had chronic musculoskeletal multi-site pain, defined as pain persisting for more than three months in at least two anatomical regions, and had not yet received a formal diagnosis. In addition to patients subsequently diagnosed with FM according to the 2016 ACR criteria, individuals with common non-inflammatory musculoskeletal pain conditions, including OA, myofascial pain, and periarticular disorders (*e.g.*, rotator cuff tendinopathy, greater trochanteric pain syndrome), were also included. The diagnosis of these conditions was based on clinical evaluation, supplemented when clinically indicated by imaging studies such as radiography, ultrasound, or magnetic resonance imaging, in order to confirm the diagnostic suspicion.

Exclusion criteria comprised a prior diagnosis of FM or any condition that could significantly interfere with clinical or clinimetric assessment. Specifically, patients with coexisting chronic inflammatory arthritis, uncontrolled malignancies, major neurological disorders (*e.g.*, dementia, Parkinson's disease, uncontrolled major depression, or severe neuropathic pain conditions such as spinal cord injury, post-herpetic neuropathy, or chemotherapy-induced neuropathy), and those unable to adequately understand Italian were excluded.

The procedures conducted in this study were carried out in accordance with the Declaration of Helsinki and approved by the ethics committee (Comitato Etico Unico Regionale [CERM], approval number 1970/AV2).

Fibromyalgia screening tools

All participants completed three validated self-administered FM screening instruments: the FiRST (25), the NFF (26), and the SIFIS (27). Questionnaires were completed before the specialist assessment to ensure independent symptom reporting.

The FiRST is a brief six-item self-report tool developed for the rapid identification of patients likely to meet FM criteria in both primary and specialty care (25). Its items cover core domains of the FM clinical phenotype, including diffuse and widespread musculoskeletal pain, involvement of multiple body regions, fatigue and morning stiffness, sleep disturbances, cognitive difficulties (fibro fog), and hypersensitivity to mechanical or thermal stimuli. Items are scored dichotomously (yes/no), and a total score ≥5 has been identified as the optimal cut-off for a positive screen. In its original validation, the FiRST showed a sensitivity of 90.5% and specificity of 85.7% *versus* the ACR 1990 criteria (15), with an AUC of 0.92, demonstrating excellent discriminative ability.

The NFF was developed to operationalise nociplastic pain features in the clinical identification of FM (26). It comprises 11 dichotomous items assessing key nociplastic characteristics such as generalised or migratory pain, non-anatomical pain distribution, morning fatigue, symptom exacerbation with stress, hyperalgesia, poor response to conventional analgesics, and affective-social contributors to pain. All items refer to symptoms experienced in the preceding three months. The total score reflects the number of affirmative responses, with ≥4 representing the optimal cut-off for distinguishing FM from other chronic pain conditions. In preliminary validation, the NFF showed a sensitivity of approximately 86%, specificity of 83%, and an AUC of 0.89 relative to expert diagnosis and ACR criteria; strong concordance with the 2011 and 2016 ACR criteria was also reported (6, 16). Cross-cultural data further support its generalisability: a Mexican validation study in FM *versus* generalised OA controls reported an AUC-ROC of 0.980, with ≥4 yielding a sensitivity of 86.49% and specificity of 96.15% (31).

The SIFIS questionnaire was developed in Italian-speaking patients through a rigorous, standardised, five-phase methodological process to ensure scientific validity and clinical practicality (27). It is specifically intended for screening individuals presenting with chronic widespread musculoskeletal pain who have not yet received a diagnosis of FM. The SIFIS includes dichotomous items addressing widespread pain distribution and somatic-functional symptoms consistent with contemporary ACR criteria. Its preliminary validation was conducted in 284 patients with multi-site pain (27), of whom 230 (80.9%) were diagnosed with FM, providing a robust cohort for assessing diagnostic performance. A threshold of ≥4 positive items corresponded to a post-test probability of FM ≥80% (range 81.8-87.7%), indicating strong diagnostic utility. Using a Bayesian Analysis Model, the instrument demonstrated a sensitivity of approximately 70% and specificity of 87%, with an AUC of 0.82. The item-specific likelihood ratios (LR+) ranged from 3.37 to 5.00. The SIFIS also ex-

hibited excellent test-retest reliability (intraclass correlation coefficient 0.96) and acceptable internal consistency (Cronbach's alpha 0.68), supporting its reproducibility and clinical applicability (27).

Statistical analysis

Descriptive statistics were used to summarise demographic and clinical characteristics. Continuous variables were reported as mean and standard deviation (SD) or as median and interquartile range (IQR) when data were not normally distributed. Categorical variables were presented as frequencies and percentages. Group differences between FM patients and controls were assessed according to the distributional properties of the data. For continuous variables meeting the assumption of normality, comparisons were performed using Student's t-test; when this assumption was violated, the non-parametric Mann-Whitney U test was applied. Categorical variables were analysed using the chi-square (χ^2) test or Fisher's exact test, as appropriate, to accommodate small sample sizes.

Receiver operating characteristic (ROC) curve analysis was performed to compare the discriminatory ability of the three screening tools, using the FM diagnosis based on the 2016 ACR criteria as the dichotomous reference standard. When multiple instruments aim to measure the same construct or screen for the same condition, ROC analysis offers a rigorous basis for comparative evaluation. The Area Under the Curve (AUC) provides a scalar summary of accuracy, with values ranging from 0.50 (chance level) to 1.00 (perfect discrimination). The optimal threshold for each questionnaire was derived using the Youden Index (J = sensitivity + specificity – 1), which identifies the point on the ROC curve that maximises the trade-off between sensitivity and specificity and corresponds to the maximum distance from the chance diagonal (32). Statistical comparisons of AUC values across questionnaires were conducted using the method described by DeLong *et al*. (33). In contrast to traditional discrimination metrics such as AUC-ROC, which assess a model's capacity to distinguish cases from non-cases independently of clinical context (28), DCA evaluates prediction models within a decision-analytic framework that explicitly incorporates threshold probability (30). Based on expected utility theory and the threshold model of clinical decision-making (34), DCA quantifies net benefit across a range of clinically meaningful probability thresholds. Net benefit reflects the balance between true-positive identifications and false-positive classifications, weighted according to the relative harm of unnecessary interventions implied by the chosen threshold (30). By comparing the net benefit of competing models with that of default strategies (*e.g.*, treating all or treating none), DCA determines whether a model offers added clinical value beyond its statistical discriminatory power (30, 35). Notably, DCA does not require explicit utility estimates, as harm–benefit trade-offs are implicitly encoded in the threshold probability. In this way, DCA complements ROC analysis by linking predictive performance to its practical implications for real-world clinical decision-making (30, 35).

All statistical analyses were performed using MedCalc® software (version 20.07; MedCalc Software, Mariakerke, Belgium).

Table I. Demographic characteristics of the studied population.

Demographic variables	FM patients (n° 106)	Controls (n° 99)	*p*
Age (mean, years ± SD)	49.5 ± 7.3	51.0 ± 6.9	n.s.
BMI (mean, kg/m² ± SD)	26.2 ± 2.6	25.5 ± 3.1	n.s.
Disease duration (mean, years ± SD)	9.9 ± 8.5	8.7 ± 8.1	n.s.
Marital status, number and (%)	80 (75.5%)	75 (75.8%)	n.s.
Educational level			
Primary school (number and percentage)	35 (33.0%)	39 (39.4%)	n.s.
Secondary school (number and percentage)	42 (39.6%)	35 (35.3%)	n.s.
High school/university (number and percentage)	29 (27.4%)	25 (25.3%)	n.s.

FM: fibromyalgia; SD: standard deviation; BMI: body mass index; n.s.: not significative.

Table II. Receiver Operating Characteristic (ROC) curve summary.

			95% CI		
Tools	AUC	Standard error	Lower	Upper	p
SIFIS	0.941	0.0143	0.913	0.969	<.001
FiRST	0.827	0.0288	0.771	0.884	<.001
NFF	0.829	0.0272	0.775	0.882	<.001

CI: confidence intervals; AUC: area under the curve; SIFIS: SImple FIbromyalgia Screening questionnaire; FiRST: Fibromyalgia Rapid Screening Tool; NFF: Nociplastic-Based Fibromyalgia Features.

Results

Demographic characteristics

A total of 205 female patients were included in the final analysis after the exclusion of 19 cases with incomplete diagnostic data. According to the 2016 ACR criteria, 106 patients (51.7%) were classified as having FM, whereas FM was ruled out in 99 patients (48.3%), who served as controls. The mean age of patients with FM was 49.5 years; 75.47% were married, and most had completed at least a secondary/high school education. On average, FM patients were moderately overweight, with a mean BMI of 26.2±2.6 kg/m². The mean duration of pain symptoms in the FM group was 9.1±1.9 years. No significant differences were observed between the two groups in terms of age, marital status, or educational level (Table I). Although the mean symptom duration was longer in the FM group (9.9±8.5 years *vs*. 8.7±8.1 years), the difference did not reach statistical significance (p=0.082).

ROC curve analysis

All three instruments demonstrated excellent discriminatory performance relative to the 2016 ACR criteria, with the highest AUC observed for the SIFIS (AUC=0.94), followed by the NFF (AUC=0.83) and the FiRST

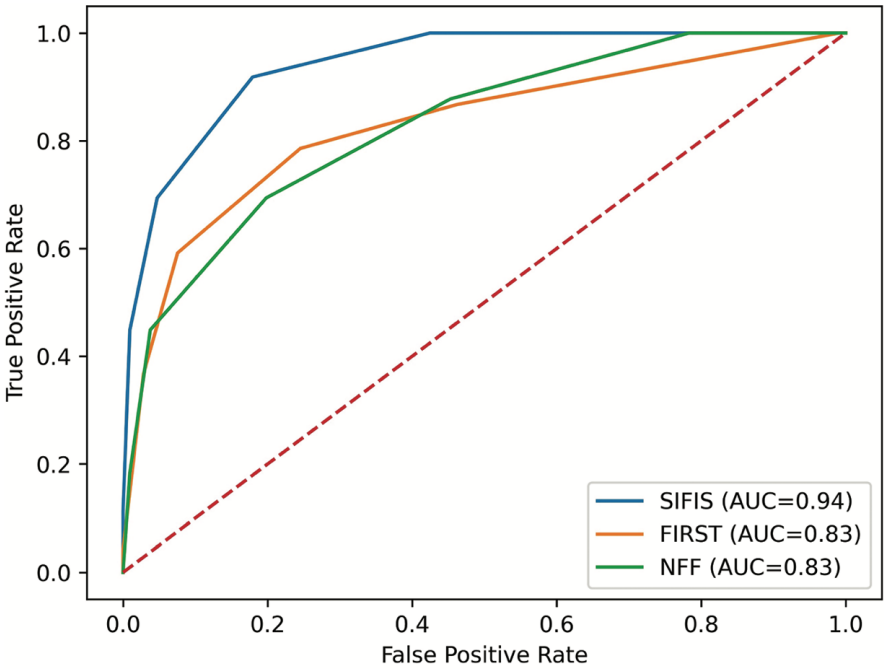

Fig. 1. Receiver Operating Characteristic (ROC) curves comparing the diagnostic performance of the SIFIS, FIRST and NFF.

(AUC=0.83) (Table II). These findings were further supported by visual inspection (Fig. 1).

A SIFIS cut-off score ≥4 provided optimal discrimination between FM and non-FM chronic pain conditions, yielding a sensitivity of 82.08%, a specificity of 91.84%, and an overall accuracy of 86.76%. A FiRST cut-off score ≥5 achieved the highest rate of correct patient classification, with a sensitivity of 75.47%, specificity of 69.39%, and accuracy of 76.96%. Using a cut-off score ≥4, the NFF showed a sensitivity of 80.19%, specificity of 69.39%, and an accuracy of 75.00% (Table III).

Pairwise DeLong test comparing AUC indicated that SIFIS performed significantly better than both FiRST (z=3.826, $p<0.001$) and NFF (z=3.688, $p<0.001$), while FiRST and NFF did not differ significantly (z=-0.026, $p=0.979$) (Table IV).

Decision curve analysis (DCA)

DCA demonstrated that SIFIS consistently provided greater net benefit, indicating a more favourable balance between true-positive identification and false-positive overclassification. Although FiRST and NFF yielded positive net benefit across portions of the threshold range, their curves remained consistently below that of SIFIS. Moreover, the confidence intervals showed limited overlap in several intermediate regions, underscoring the robustness of SIFIS's comparative advantage (Fig. 2).

Discussion

The present study examined the diagnostic performance and clinical utility of three brief self-administered screening tools for FM, SIFIS, FiRST and NFF, in a cohort of patients with chronic pain. By integrating ROC analysis with DCA, the study assessed not only the discriminative accuracy of these instruments but also their potential impact on clinical decision-making. The findings indicate that the SIFIS questionnaire demonstrated superior diagnostic accuracy and provided the greatest net clinical benefit across a wide range of probability thresholds, suggesting that it may serve as a particularly effective screening tool for identifying FM in routine clinical practice.

FM is a multifaceted nociplastic pain condition characterized by widespread musculoskeletal pain accompanied by fatigue, sleep disturbances, cognitive impairment, and psychological symptoms (5). Despite growing recognition of its underlying pathophysiology, FM remains challenging to diagnose, as no specific laboratory or imaging biomarkers are currently available and symptoms frequently overlap with those of other chronic pain disorders (2, 6, 20). Consequently, diagnostic delays are common, and patients often experience prolonged periods of uncertainty before receiving an appropriate diagnosis (11, 12, 21). In this context, the use of rapid

Table III. Diagnostic accuracy of SIFIS, FiRST and NFF.

	Diagnostic accuracy - SIFIS			Diagnostic accuracy - FiRST			Diagnostic accuracy - NFF		
		95% CI			95% CI			95% CI	
	Result	Lower	Upper	Result	Lower	Upper	Result	Lower	Upper
Sensitivity	82.08 %	73.43 %	88.85 %	75.47 %	66.16 %	83.31 %	80.19 %	71.32 %	87.30 %
Specificity	91.84 %	84.55 %	96.41 %	78.57 %	69.13 %	86.22 %	69.39 %	59.26 %	78.30 %
LR+	10.05	5.15	19.65	3.52	2.37	5.23	2.62	1.92	3.58
LR-	0.20	0.13	0.30	0.31	0.22	0.44	0.29	0.19	0.43
PPV	91.58 %	84.77 %	95.51 %	79.21 %	71.97 %	84.96 %	73.91 %	67.45 %	79.48 %
NPV	82.57 %	75.84 %	87.73 %	74.76 %	67.62 %	80.77 %	76.40 %	68.35 %	82.92 %
Accuracy	86.76 %	81.33 %	91.09 %	76.96 %	70.57 %	82.55 %	75.00 %	68.47 %	80.78 %

SIFIS: SImple FIbromyalgia Screening questionnaire; FiRST: Fibromyalgia Rapid Screening Tool; NFF: Nociplastic-Based Fibromyalgia Features; CI: confidence intervals; LR+: positive likelihood ratio; LR-: negative likelihood ratio; PPV: positive predictive value; NPV: negative predictive value.

Table IV. Pairwise AUC comparisons.

		95% CI			
Comparisons	AUC difference	Lower	Upper	z	p
SIFIS vs. FIRST	0.11403	0.0556	0.1724	3.826	<0.001
SIFIS vs. NFF	0.11287	0.0521	0.1737	3.638	<0.001
FIRST vs. NFF	-0.00116	-0.0858	0.0835	-0.026	0.979

CI: confidence intervals; AUC: area under the curve; SIFIS: SImple FIbromyalgia Screening questionnaire; FiRST: Fibromyalgia Rapid Screening Tool; NFF: Nociplastic-Based Fibromyalgia Features.

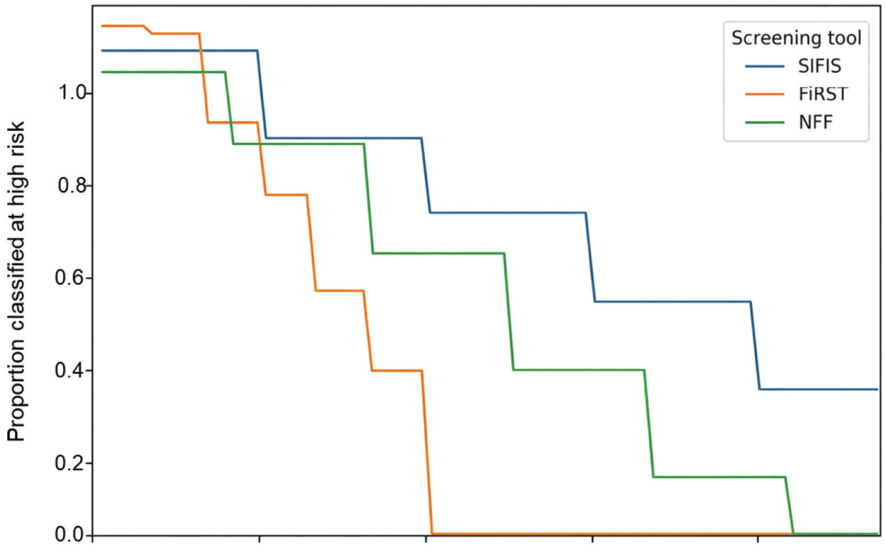

Fig. 2. Decision Curve Analysis (DCA) showing comparing the net benefits of SIFIS, FiRST and NFF across a broad range of clinically relevant threshold probabilities.

screening instruments has been proposed as a pragmatic strategy to support the early identification of patients with FM-related features, particularly in non-specialist settings such as primary care or general outpatient clinics.

Previous research on FM screening instruments has reported moderate-to-high diagnostic performance, with AUC values typically ranging from approximately 0.75 to 0.90 depending on the study population (16, 25). The FiRST questionnaire, for instance, has been validated as a rapid screening tool with good sensitivity for identifying FM in rheumatology settings; however, its performance may be less robust in heterogeneous chronic pain populations (25). Similarly, the NFF was developed to capture clinical features associated with nociplastic pain mechanisms, but its diagnostic utility may vary depending on clinical context and patient characteristics (36). The superior accuracy of SIFIS observed in the present study may reflect its comprehensive symptom coverage and its specific design to capture clinical domains particularly relevant to FM.

A notable strength of this study is the integration of DCA, which enabled evaluation of the clinical usefulness of the screening tools beyond conventional diagnostic metrics. Traditional indicators such as sensitivity, specificity, and AUC are essential for assessing test performance but do not directly address the clinical consequences of employing a diagnostic tool in real-world practice (37). DCA provides a complementary framework by estimating the net benefit of a diagnostic strategy across a range of clinically relevant threshold probabilities, thereby incorporating the balance between potential benefits and harms of diagnostic decisions (30). In the present analysis, SIFIS consistently conferred the highest net benefit across a wide spectrum of thresholds.

Early recognition of FM carries meaningful clinical implications. Delayed diagnosis has been associated with increased healthcare utilization, repeated diagnostic investigations, and reduced quality of life (22). Evidence indicates that individuals with FM often experience diagnostic delays lasting several years, frequently consulting multiple healthcare providers before receiving an accurate diagnosis (12, 38). Implementing effective screening strategies in routine clinical practice may therefore facilitate earlier identification, timelier specialist referral, and prompt initiation of appropriate management pathways. Moreover, providing patients with an explanatory framework for their symptoms can have a substantial psychological impact, reducing uncertainty and enhancing engagement with multimodal treatment strategies (39).

An additional strength of this study is the inclusion of a clinically relevant control group comprising patients with common non-inflammatory musculoskeletal disorders such as OA, mechanical spinal disorders, and periarticular syndromes. These conditions frequently coexist with chronic pain and may share certain clinical features with FM, complicating differential diagnosis in everyday practice (13, 14). The ability of SIFIS to maintain high diagnostic accuracy in this heterogeneous context further supports its potential utility as a screening tool in real-world clinical environments.

Nevertheless, several limitations should be acknowledged. First, the study population was composed exclusively of female patients. Although FM is more prevalent among women, this limits the generalisability of the findings to male populations (40, 41). The application of this type of screening tools has been scarcely investigated in male subjects presenting with a clinical picture of chronic multisite pain. Given the lower prevalence of FM in men, data derived from multicentre cohorts are required to achieve adequate validation and meaningful comparisons among these instruments within the male population. Second, the study was conducted in a single tertiary-care centre, introducing potential referral bias and limiting

external validity. Additional validation studies in broader clinical settings, including primary care and community-based cohorts, are needed to confirm the robustness of these findings. Future research should also explore the integration of screening tools such as SIFIS with emerging phenotypic approaches designed to identify nociplastic pain mechanisms, which may further refine diagnostic pathways and support personalized management strategies. Third, no formal Italian-language validation of the FiRST and NFF was performed. The two questionnaires were simply translated and administered to patients, and the Italian versions (Supplementary Material) should undergo a more rigorous and formal validation process. This aspect may limit the comparability of the results with those of future studies conducted in Italian-speaking patients. Moreover, the SIFIS was originally developed as an Italian-language questionnaire, which may account for a degree of methodological asymmetry, as it did not undergo a translation process from English into Italian. Finally, the high prevalence of FM in the studied population (51.7%) should be taken into consideration. This may represent a selection bias, given that the setting is a tertiary-level rheumatology centre, potentially limiting the external validity of the findings in primary care settings.

In conclusion, this study demonstrates that the SIFIS questionnaire exhibits excellent diagnostic performance and superior clinical utility compared with FiRST and NFF in patients presenting with chronic pain. This superiority will need to be demonstrated in broader settings.

The adoption of efficient screening tools may facilitate earlier recognition of FM and contribute to improved patient care.

References

1. BENNETT R: Fibromyalgia: present to future. *Curr Pain Headache Rep* 2004; 8(5): 379-84. https://doi.org/10.1007/s11916-996-0011-5
2. CLAUW DJ: Fibromyalgia: a clinical review. *JAMA* 2014; 311(15): 1547-55. https://doi.org/10.1001/jama.2014.3266
3. 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. https://doi.org/10.1186/1472-6963-10-102
4. CLAUW DJ, D'ARCY Y, GEBKE K, SEMEL D, PAUER L, JONES KD: Normalizing fibromyalgia as a chronic illness. *Postgrad Med* 2018; 130(1): 9-18. https://doi.org/10.1080/00325481.2018.1411743
5. BIDARI A, GHAVIDEL-PARSA B: Nociplastic pain concept, a mechanistic basis for pragmatic approach to fibromyalgia. *Clin Rheumatol* 2022; 41(10): 2939-47. https://doi.org/10.1007/s10067-022-06229-5
6. WOLFE F, CLAUW DJ, FITZCHARLES MA *et al*.: 2016 Revisions to the 2010/2011 fibromyalgia diagnostic criteria. *Semin Arthritis Rheum* 2016; 46(3): 319-29. https://doi.org/10.1016/j.semarthrit.2016.08.012
7. BRANCO JC, BANNWARTH B, FAILDE I *et al*.: Prevalence of fibromyalgia: a survey in five European countries. *Semin Arthritis Rheum* 2010; 39(6): 448-53. https://doi.org/10.1016/j.semarthrit.2008.12.003
8. CABO-MESEGUER A, CERDÁ-OLMEDO G, TRILLO-MATA JL: Fibromyalgia: Prevalence, epidemiologic profiles and economic costs. *Med Clin* (Barc) 2017; 149(10): 441-8. https://doi.org/10.1016/j.medcli.2017.06.008
9. RUIZ-MONTERO PJ, SEGURA-JIMENEZ V, ALVAREZ-GALLARDO IC *et al*.: Fibromyalgia Impact Score in Women with Fibromyalgia Across Southern, Central, and Northern Areas of Europe. *Pain Physician* 2019; 22(5): E511-E516.
10. SALAFFI F, DE ANGELIS R, GRASSI W *et al*.: 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(6): 819-28.
11. ALI S, BANDA A, RAMIREZ-MATOS C, COHEN P: From Years of Pain to a Name: Fibromyalgia Uncovered. *Cureus* 2025; 17(12): e98981. https://doi.org/10.7759/cureus.98981
12. SALAFFI F, FARAH S, BIANCHI B, LOMMANO MG, DI CARLO M: Delay in fibromyalgia diagnosis and its impact on the severity and outcome: a large cohort study. *Clin Exp Rheumatol* 2024; 42(6): 1198-1204. https://doi.org/10.55563/clinexprheumatol/ta9xtc
13. WOLFE F, WALITT BT, KATZ RS, LEE YC, MICHAUD KD, HÄUSER W: Longitudinal patterns of analgesic and central acting drug use and associated effectiveness in fibromyalgia. *Eur J Pain* 2013; 17(4): 581-6. https://doi.org/10.1002/j.1532-2149.2012.00234.x
14. DI FRANCO M, IANNUCCELLI C, BAZZICHI L *et al*.: Misdiagnosis in fibromyalgia: a multicentre study. *Clin Exp Rheumatol* 2011; 29 (Suppl. 69): S104-108.
15. 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(2): 160-72. https://doi.org/10.1002/art.1780330203
16. 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(5): 600-10. https://doi.org/10.1002/acr.20140
17. LEBEAU FOUSTOUKOS T, LECLERCQ I, BLANCHARD M, HÜGLE T: Evidence-Based Self-Management Strategies for Fibromyalgia: Foundations for Digital Therapeutic Applications. Interact *J Med Res* 2026; 15: e67523. https://doi.org/10.2196/67523
18. COHEN-BITON L, BUSKILA D, NISSAN-HOLTZ-GANNOT R: Review of Fibromyalgia (FM) Syndrome Treatments. *Int J Environ Res Public Health* 2022; 19(19): 12106. https://doi.org/10.3390/ijerph191912106
19. HERNANDO-GARIJO I, JIMÉNEZ-DEL-BARRIO S, MINGO-GÓMEZ T, MEDRANO-DE-LA-FUENTE R, CEBALLOS-LAITA L: Effectiveness of non-pharmacological conservative therapies in adults with fibromyalgia: A systematic review of high-quality clinical trials. *J Back Musculoskelet Rehabil* 2022; 35(1): 3-20. https://doi.org/10.3233/BMR-200282
20. HÄUSER W, ABLIN J, FITZCHARLES MA *et al*.: Fibromyalgia. *Nat Rev Dis Primers* 2015; 1: 15022. https://doi.org/10.1038/nrdp.2015.22
21. ARNOLD LM, BENNETT RM, CROFFORD LJ *et al*.: AAPT Diagnostic Criteria for Fibromyalgia. *J Pain* 2019; 20(6): 611-28. https://doi.org/10.1016/j.jpain.2018.10.008
22. MACFARLANE GJ, KRONISCH C, DEAN LE *et al*.: EULAR revised recommendations for the management of fibromyalgia. *Ann Rheum Dis* 2017; 76(2): 318-28. https://doi.org/10.1136/annrheumdis-2016-209724
23. HERMAN AM, BERRYMAN C, STANTON TR: Heightened protective decision-making related to physical, but not cognitive, effort in individuals with fibromyalgia. *J Pain* 2025; 34: 105504. https://doi.org/10.1016/j.jpain.2025.105504
24. KÜLEKÇIOĞLU S: Diagnostic difficulty, delayed diagnosis, and increased tendencies of surgical treatment in fibromyalgia syndrome. *Clin Rheumatol* 2022; 41(3): 831-7. https://doi.org/10.1007/s10067-021-05970-7
25. PERROT S, BOUHASSIRA D, FERMANIAN J *et al*.: Development and validation of the Fibromyalgia Rapid Screening Tool (FiRST). *Pain* 2010; 150(2): 250-6. https://doi.org/10.1016/j.pain.2010.03.034
26. GHAVIDEL-PARSA B, BIDARI A, ATRKAR-ROUSHAN Z, KHOSOUSI MJ: Implication of the Nociplastic Features for Clinical Diagnosis of Fibromyalgia: Development of the Preliminary Nociplastic-Based Fibromyalgia Features (NFF) Tool. *ACR Open Rheumatol* 2022; 4(3): 260-8. https://doi.org/10.1002/acr2.11390
27. SALAFFI F, FARAH S, BECI G, SCHETTINO M, CAROTTI M, DI CARLO M: Development and validation of the SImple FIbromyalgia Screening questionnaire for improving the recognition of fibromyalgia in daily practice. *Clin Exp Rheumatol* 2020; 38 (Suppl. 123): S9-16.
28. HANLEY JA, MCNEIL BJ: The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 1982; 143(1): 29-36. https://doi.org/10.1148/radiology.143.1.7063747
29. ZWEIG MH, CAMPBELL G: Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. *Clin Chem* 1993; 39(4): 561-77. Erratum in: Clin Chem 1993; 39(8): 1589.

30. VICKERS AJ, ELKIN EB: Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Making* 2006; 26(6): 565-74. https://doi.org/10.1177/0272989X06295361
31. SALDANA-CAMPOS EM, NOVOA-ISLAS M, CENICEROS-RUIZ B *et al*.: Cross-cultural adaptation and validation of the Nociplastic-based Fibromyalgia Features (NFF) diagnostic tool in Mexican fibromyalgia patients. *Ann Rheum Dis* 2024; 83 (Suppl. 1): 1314.1. https://doi.org/10.1136/annrheumdis-2024-eular.6229
32. YOUDEN WJ: Index for rating diagnostic tests. *Cancer* 1950; 3(1): 32-5. https://doi.org/10.1002/1097-0142(1950)3:1<32::aid-cncr2820030106>3.0.co;2-3
33. DELONG ER, DELONG DM, CLARKE-PEARSON DL: Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 1988; 44(3): 837-45.
34. PAUKER SG, KASSIRER JP: The threshold approach to clinical decision making. *N Engl J Med* 1980; 302(20): 1109-17. https://doi.org/10.1056/NEJM198005153022003
35. VICKERS AJ, VAN CALSTER B, STEYERBERG EW: Net benefit approaches to the evaluation of prediction models, molecular markers, and diagnostic tests. *BMJ* 2016; 352: i6. https://doi.org/10.1136/bmj.i6
36. KOSEK E, COHEN M, BARON R *et al*.: Do we need a third mechanistic descriptor for chronic pain states? *Pain* 2016; 157(7): 1382-6. https://doi.org/10.1097/j.pain.0000000000000507
37. HAJIAN-TILAKI K: Receiver Operating Characteristic (ROC) Curve Analysis for Medical Diagnostic Test Evaluation. *Caspian J Intern Med* 2013; 4(2): 627-35.
38. HÄUSER W, PERROT S, SOMMER C, SHIR Y, FITZCHARLES MA: Diagnostic confounders of chronic widespread pain: not always fibromyalgia. *Pain Rep* 2017; 2(3): e598. https://doi.org/10.1097/PR9.0000000000000598
39. SARACOGLU I, AKIN E, AYDIN DINCER GB: Efficacy of adding pain neuroscience education to a multimodal treatment in fibromyalgia: A systematic review and meta-analysis. *Int J Rheum Dis* 2022; 25(4): 394-404. https://doi.org/10.1111/1756-185X.14293
40. QUEIROZ LP: Worldwide epidemiology of fibromyalgia. *Curr Pain Headache Rep* 2013; 17(8): 356. https://doi.org/10.1007/s11916-013-0356-5
41. IANNUCCELLI C, FAVRETTI M, DOLCINI G *et al*.: The impact of sex and gender on Fibromyalgia Syndrome: data from the Italian Fibromyalgia Registry. *Intern Emerg Med* 2026. https://doi.org/10.1007/s11739-026-04277-2