Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Development and validation of MIGRAGILE: a brief digital patient-reported outcome measure to identify disabling migraine in primary care and support referral to specialised care

Teresa Marco Galindo
mtmarco@clinic.cat
Hospital Clínic de Barcelona
Víctor Obach
Hospital Clínic de Barcelona
Neus Fabregat Fabra
Hospital Clínic de Barcelona
Anna Grajea Ilaria
Hospital Clínic de Barcelona
Ángeles Sosa
August Pi i Sunyer Biomedical Research Institute
Adelaida Zabalegui
Hospital Clínic de Barcelona
Santiago Fernández Fernández
Hospital Clínic de Barcelona

Research Article

Keywords: patient-reported outcome measure, migraine, primary care, clinical stratification, referral decision-making, validation

Posted Date: May 20th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9425208/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: Competing interest reported. TMG has received honoraria for participating in medical education activities from Eli Lilly, Organon, Novartis, Teva and Lundbeck. TMG has also received

consulting fees from AbbVie and Lundbeck. In addition, TMG has received competitive public grants to fund nurse intensification contracts in the Strategic Plan for Health Research and Innovation 2022–2027 of the Government of Catalonia. VOB has received honoraria for participating in medical education activities from AbbVie, Eli Lilly, Organon, Novartis, Teva, Lundbeck and Dr. Reddy's. SFF has received honoraria for participating in medical education activities from Pfizer, Lundbeck and Organon, NFF has received honoraria for participating in medical education activities from Eli Lilly, Novartis, Teva, Lundbeck, Pfizer.

Abstract

Background

Migraine is a highly prevalent neurological disorder and a leading cause of disability worldwide. Despite the availability of effective preventive therapies, many individuals with disabling migraine remain underdiagnosed or undertreated, particularly in primary care. Patient-reported outcome measures (PROMs) can help capture disease burden directly from patients and support early clinical stratification. We aimed to develop and validate MIGRAGILE, a brief 11-item digital self-administered PROM designed to identify, in primary care, patients who are likely to have disabling migraine and to support structured referral decisions to specialised care.

Methods

We conducted a multicentre study across 20 primary care centres and one tertiary hospital in Spain. MIGRAGILE was developed through systematic literature review, multidisciplinary expert consensus, and direct patient input during development and face-validity assessment, followed by pilot testing and psychometric validation. Content validity was assessed using the Hambleton-Rovinelli item-domain congruence procedure, and face validity by patients and clinicians. Test-retest reliability (7–14 days) and inter-observer agreement (three neurologists, two advanced practice nurses, one clinical researcher) were evaluated using weighted Cohen's kappa, Fleiss' kappa, and intraclass correlation coefficients (ICC). Criterion validity/classification performance was assessed against a headache-specialist ICHD-3 reference diagnosis, blinded to MIGRAGILE results; the optimal cut-off was defined using the Youden index. Convergent validity was examined using HIT-6 and MSQ v2.1, and associations with care pathways and first-visit management decisions were analysed.

Results

Of 300 participants, 277 had complete data for the classification performance analysis. MIGRAGILE showed high classification performance for disabling migraine (high-frequency episodic migraine/chronic migraine), with sensitivity 93.7% (95% CI 91.7–95.3) and specificity 94.0% (95% CI 90.2–96.8); AUC was 0.91 (95% CI 0.87–0.95). Test-retest reliability was excellent for quantitative items (ICC 0.998–0.999), and categorical items showed good-to-excellent stability (weighted κ 0.784–1.000). Overall multi-rater agreement for referral categories was moderate-to-substantial (Fleiss' κ 0.668), and was highest for preferential headache-unit referral (κ 0.820). MIGRAGILE classification was strongly associated with HIT-6 (χ^2=115.99, $p < 0.001$; Cramér's V = 0.46), MSQ (H = 45.62, $p < 0.001$; ε^2=0.159), and care pathways and treatment decisions (χ^2=234.8, $p < 0.001$; Cramér's V = 0.65).

Conclusions

MIGRAGILE is a brief, digital and psychometrically robust PROM that accurately identifies patients with disabling migraine and supports clinically meaningful stratification in primary care. Its association with patient-reported burden and real-world clinical decisions suggests that it may facilitate more appropriate referral to specialised care and contribute to more timely and targeted management in routine practice.

INTRODUCTION

Migraine is a highly prevalent neurological disorder and one of the leading causes of disability worldwide, affecting approximately 12% of the general population [1]. Its impact extends beyond the frequency of headache attacks and substantially impairs physical, emotional and social functioning [2–5]. Despite this burden, migraine has historically been underdiagnosed and undertreated, partly due to the social normalisation of the condition and the limited availability of migraine-specific preventive therapies in the past, which contributed to therapeutic inertia in its management [6, 7].

In recent years, major advances in acute and preventive migraine treatment have improved patients' health-related quality of life (HRQoL) [8, 9]. Nevertheless, a substantial proportion of patients still do not receive adequate care, as shown by the OVERCOME study [6, 7, 10]. Poor disease control is associated with an increased risk of chronification and medication overuse, underscoring the need to optimise early identification and management [9].

The introduction of therapies targeting the calcitonin gene-related peptide (CGRP) pathway, alongside established options such as onabotulinumtoxinA, has shifted the therapeutic landscape towards more personalised care models [8, 9]. This has increased the importance of early identification of patients with higher disease burden who are more likely to benefit from specialist input, and of structured, patient-centred assessment to optimise care pathways and outcomes [11].

In this context, patient-reported outcome measures (PROMs) play an important role by providing direct information on symptoms, functioning and disease impact from the patient perspective, which may not be fully captured through conventional clinical assessment [11, 12]. Several self-administered tools have been developed for migraine screening or clinical characterisation. Instruments such as ID-Migraine™, MS-Q and HARDSHIP [13–15] have demonstrated utility for case-finding, particularly in epidemiological contexts, while tools such as VARS and MHQ-8 [16, 17] assess specific clinical domains including aura and sensory hypersensitivity. However, most available tools were developed primarily for case-finding or clinical characterisation and were not designed to identify disabling migraine or to support actionable stratification and referral decisions in primary care. In particular, they do not simultaneously integrate key dimensions—attack frequency, associated disability and functional impact—into an actionable classification to inform care. A summary comparison of existing self-administered tools and MIGRAGILE is provided in Supplementary Tables S5a–S5b.

Against this backdrop—characterised by the availability of targeted preventive therapies and the need to optimise access to specialist resources—there is a growing need for patient-reported instruments that

not only identify migraine, but also stratify clinical severity and support structured decision-making within care pathways [7–9, 12, 18].

The aim of this study was to develop and validate MIGRAGILE, a novel brief digital PROM designed to assess key clinical dimensions of migraine (attack frequency, disability burden, core migraine features, and acute/preventive treatment patterns) from the patient perspective, to identify, in primary care, patients who are likely to have disabling migraine, and to support structured referral decisions to specialised care.

METHODS

Ethics approval

The study was conducted in accordance with Good Clinical Practice (GCP), the Declaration of Helsinki, and the EU General Data Protection Regulation (GDPR 2016/679). Ethical approval was obtained from the Clinical Research Ethics Committee of Hospital Clínic de Barcelona (reference number: HCB/2023/0606). All participants provided written informed consent prior to participation.

Study design, settings and participants

A multicentre, cross-sectional study with an embedded implementation component was conducted between June 2023 and May 2025 across 20 primary care teams and their tertiary referral hospital in Spain. Detailed information on participating centres is available from the authors upon reasonable request. The study was considered interventional insofar as MIGRAGILE was integrated into routine clinical workflow and used to inform referral decisions according to a pre-specified protocol. MIGRAGILE captures four prespecified clinical domains—headache frequency and disability, core migraine features/red flags, acute treatment behaviour, and preventive treatment history—and generates two clinically actionable outputs: reason for consultation and recommended referral level/care pathway.

Workflow and activation of MIGRAGILE

Patients consulting their general practitioner (GP) in primary care for headache were informed about MIGRAGILE after the routine clinical consultation and were offered the option to complete the questionnaire to support subsequent assessment by the tertiary headache unit. Patients accessed MIGRAGILE via a QR code and completed the questionnaire remotely at home in a self-administered format. Responses were captured and stored in REDCap. GPs did not have access to MIGRAGILE results. An advanced practice nurse (APN) within the tertiary headache unit reviewed each completed questionnaire and applied the pre-specified interpretation rules to generate the reason for consultation and recommended referral level/care pathway. APN-derived outputs were generated prior to, and independently of, the headache-specialist reference diagnosis.

Consecutive recruitment was used to minimise selection bias. Eligible participants were adults (≥ 18 years) attending primary care for headache evaluation who were able to provide informed consent and complete a digital questionnaire in Spanish (Spain). Exclusion criteria were confirmed secondary headache, severe cognitive impairment, hospitalisation at the time of assessment, or substantial language barriers.

Three independent samples were used: (1) a pilot usability sample (n = 70); (2) a test–retest reliability subsample (n = 70); and (3) the validation cohort (n = 300). The test–retest interval of 7–14 days was selected in accordance with methodological recommendations for temporal stability studies, balancing the likelihood of true clinical change against recall effects. Sample sizes followed published recommendations for PROM development studies [19–21].

Instrument development

Questionnaire development followed a structured process consistent with international methodological standards for PRO instruments: (1) conceptual framework definition, (2) item generation, (3) content validation, and (4) pilot testing.

A multidisciplinary panel comprising three headache neurologists, two advanced practice nurses specialising in headache, one expert patient and one health outcomes researcher generated items based on a systematic literature review (PubMed, Scopus and Embase) and ICHD-3 diagnostic criteria. The full patient version of MIGRAGILE is provided in Supplementary Files S1–S2, and the technical description of items, domains, response formats and variable types is presented in Supplementary Table S1.

Content validity was assessed using the Hambleton–Rovinelli item–domain congruence procedure [22]. Items were phrased affirmatively and designed to capture a single concept to facilitate completion. Experts rated relevance and clarity using 4-point Likert scales, and unanimous agreement was required for final item inclusion.

Pilot testing

The preliminary version of MIGRAGILE was self-administered digitally by 70 participants to assess comprehensibility, conceptual relevance, semantic clarity and usability. Completion time, item-level missingness and qualitative feedback were recorded. Usability was evaluated using the System Usability Scale (SUS), a widely validated instrument [23]. Items and response options are detailed in Supplementary Table S1 and Supplementary Files S1–S2.

Psychometric evaluation

Psychometric properties were evaluated in accordance with internationally accepted methodological standards for validating PRO instruments.

Face validity.

Face validity was assessed through independent review by patients and clinicians, focusing on comprehensibility, semantic clarity, clinical relevance and ease of response.

Test–retest reliability.

Temporal stability was evaluated separately for categorical and quantitative variables. Weighted Cohen's κ coefficients were calculated for categorical variables [24]. For quantitative variables, intraclass correlation coefficients (ICC) were estimated using a two-way random-effects model with absolute agreement (ICC [2, 1]) [25]. κ values were interpreted according to Landis and Koch, and ICC values according to Koo and Li. Item-level test–retest results are provided in Supplementary Table S2.

Inter-observer agreement.

Inter-observer agreement in the clinical interpretation of MIGRAGILE outputs was evaluated by six independent raters (three headache neurologists, two advanced practice nurses and one medical clinical researcher) who assessed the same cases using two complementary approaches. First, pairwise weighted Cohen's κ coefficients were calculated for all possible rater combinations [26–28]. Second, overall agreement across multiple raters was estimated using Fleiss' κ, analysed separately for each referral category [28–30]. Patterns of agreement were explored using hierarchical clustering based on a dissimilarity matrix defined as (1 − κ). The full pairwise agreement matrix for reason for consultation is provided in Supplementary Table S3. Fleiss' κ values by referral category (with Z statistics and p values) are presented in Supplementary Table S4.

Criterion validity / classification performance.

Criterion validity and classification performance were evaluated in the validation sample prior to systematic implementation of the care algorithm. MIGRAGILE classification was compared with the clinical reference diagnosis established by headache-specialist neurologists according to ICHD-3 criteria, considered the gold standard in the absence of objective biomarkers. Neurologists were blinded to MIGRAGILE results. Sensitivity, specificity, predictive values, likelihood ratios and the area under the ROC curve (AUC) were calculated [31, 32]. The optimal cut-off was determined using the Youden index [33]. Ninety-five per cent confidence intervals were estimated for all classification performance indices [31, 34].

Convergent validity.

Convergent validity was explored by comparing MIGRAGILE clinical categories with validated PROMs (HIT-6 and MSQ) administered at the first visit to the tertiary headache unit [35, 36]. HIT-6 categories were compared using χ^2 tests and effect size was quantified using Cramér's V [37, 38]. MSQ scores were compared using Kruskal–Wallis tests and effect size was estimated using ε^2 [39, 40].

Clinical implementation

For the clinical implementation phase, a structured care protocol was established based on MIGRAGILE outputs to standardise referral decisions and initial management.

The referral algorithm was defined as follows:

- ≥ 8 monthly headache days with migraine features and ≥ 8 disability days: preferential referral to the headache unit (< 2 weeks).
- ≥ 8 monthly headache days with migraine features and < 8 disability days: routine referral to the headache unit (< 4 weeks).
- < 8 monthly headache days with migraine features and < 8 disability days: Telematic therapeutic education and follow-up in primary care.

Clinical decisions were made in accordance with this pre-specified protocol integrated into routine care.

Statistical analysis

Appropriate descriptive and inferential statistics were applied according to variable type and distribution. Normality of continuous variables was assessed using the Shapiro–Wilk test [41]. All tests were two-sided with a significance threshold of $p < 0.05$. Missing data were summarised descriptively and were not imputed due to the low proportion of missing values. All analyses were performed using SPSS version 29 (IBM Corp., Armonk, NY, USA). The manuscript was prepared in accordance with the COSMIN Reporting Guideline for studies on measurement properties of PROMs [42] (Supplementary File S3).

RESULTS

Sample characteristics

Table 1 summarises the sociodemographic and clinical characteristics of the validation, pilot and test–retest samples. In the validation sample (n = 300), 83.3% of participants were women and the mean age was 40.5 years (SD 13.8). The mean number of monthly headache days was 10.5 (SD 8.2). The diagnostic spectrum included low-frequency episodic migraine (≤ 7 days/month), high-frequency episodic migraine (8–14 days/month), chronic migraine (≥ 15 days/month), and other headache disorders. The pilot (n = 70) and test–retest (n = 70) samples showed comparable characteristics. No statistically significant differences were observed among the three samples in any of the sociodemographic or clinical variables analysed (all $p > 0.05$), indicating good comparability across groups. Of the 300 participants in the validation sample, 277 had complete questionnaire data and were included in the analyses. The proportion of missing data was low and, given its limited impact on the primary variables, no imputation was performed.

Table 1
Baseline sociodemographic, diagnostic and clinical characteristics of study samples.

Variable	Validation sample (n = 300)	Pilot (n = 70)	Test–retest (n = 70)	p-value
Age, mean (SD)	40.5 (13.8)	43.9 (13.0)	41.6 (12.9)	0.08
Female, n (%)	250 (83.3)	58 (83.8)	58 (83.5)	0.99
Monthly headache days (MHD), mean (SD)	10.5 (8.2)	11.4 (7.2)	11.5 (8.8)	0.62
Headache Type: n (%)				0.81
Low-frequency episodic migraine (≤ 7 MHD)	160 (53.3)	35 (50.0)	36 (51.4)	
High-frequency episodic migraine (8–14 MHD)	58 (19.4)	17 (24.3)	16 (22.8)	
Chronic migraine (≥ 15 MHD)	48 (16.0)	11 (15.7)	9 (12.8)	
Other headaches	34 (11.3)	7 (10.0)	9 (12.8)	
Aura, yes, n (%)	110 (36.6)	26 (37.1)	24 (34.4)	0.90
Pain intensity, n				0.65
mild	19	4	6	
moderate	136	33	27	
severe	145	33	37	
Educational level, n (%)				0.41
Primary education or less	27 (9.0)	8 (11.3)	9 (12.8)	
Secondary education	188 (62.7)	42 (60.0)	38 (54.2)	
University education	74 (24.7)	20 (28.5)	23 (32.8)	
Employment status, n (%)				0.36
Unemployed	38 (12.7)	11 (15.7)	10 (14.2)	
Retired	29 (9.7)	7 (10.0)	8 (11.3)	
Student	9 (3.0)	3 (4.2)	2 (2.8)	
Part-time employment	38 (12.7)	11 (15.7)	15 (21.4)	
Full-time employment	170 (56.7)	35 (50.0)	34 (48.5)	

Notes: MHD: Migraine Headache Days. SD: Standard Deviation. Statistical analysis: Continuous variables were compared using one-way ANOVA, and categorical variables using the chi-square test. A p-value < 0.05 was considered statistically significant.

Variable	Validation sample (n = 300)	Pilot (n = 70)	Test–retest (n = 70)	p-value
Other	9 (3.0)	3 (4.2)	1 (1.4)	

Notes: MHD: Migraine Headache Days. SD: Standard Deviation. Statistical analysis: Continuous variables were compared using one-way ANOVA, and categorical variables using the chi-square test. A p-value < 0.05 was considered statistically significant.

Content validity

Content validity analysis using the Hambleton–Rovinelli procedure confirmed that all 11 items were appropriately aligned with the predefined conceptual domains. Following the first Delphi round, three items (monthly headache days, acute medication use, and pain intensity) were reworded to improve semantic clarity, particularly for patients with lower health literacy. In the second round, all items achieved maximal I-CVI values for both relevance and clarity, and the overall scale reached an S-CVI/Ave of 1.00, indicating unanimous expert agreement.

Face validity

Face validity was evaluated in an independent group of patients (n = 20). Participants considered the questionnaire easy to understand, appropriately brief, and visually accessible in its digital format. No items were identified as confusing or redundant, and no additional domains were suggested.

Pilot testing and usability

In the pilot phase (n = 70), the mean System Usability Scale (SUS) score was 92.5 (SD 8.3), ranging from 75 to 100. The 25th, 50th and 75th percentiles were 87.5, 92.0 and 98.0, respectively. The mean questionnaire completion time was 5.3 minutes (SD 1.7). No item exceeded 5% missing responses. Qualitative feedback was limited to minor wording suggestions, which were incorporated without altering the clinical content of the instrument.

Test–retest reliability

Test–retest reliability was evaluated in 70 participants who completed MIGRAGILE twice, 7–14 days apart, without clinically relevant changes between administrations. Categorical items showed weighted Cohen's κ values ranging from 0.784 to 1.000 (item-level results are provided in Supplementary Table S2). For the two quantitative items, temporal stability was excellent: the intraclass correlation coefficient ICC (2,1) was 0.998 for NI1 (monthly headache days) and 0.999 for NI2 (monthly disability days) (Table 2).

Table 2
Test–retest reliability of MIGRAGILE items.

Item	Type of variable	Statistic	Value	Interpretation
Categorical items	Categorical	Weighted Cohen's κ	0.784–1.000	Good–Excellent
NI1 – MHD	Numerical	ICC (2,1)	0.998	Excellent
NI2 – DIS	Numerical	ICC (2,1)	0.999	Excellent

Notes: ICC (2,1): two-way random effects model, absolute agreement. NI1-MHD: Numeric item 1- monthly headache days. NI2-DIS: Numeric item 2- monthly disability days

Inter-observer agreement

Inter-observer agreement was assessed among six independent raters (three neurologists, two headache advanced practice nurses, and one medical clinical researcher) for two questionnaire-derived outputs: reason for consultation and referral level.

For reason for consultation, pairwise weighted Cohen's κ coefficients ranged from 0.708 to 0.878 (Supplementary Table S3). The highest agreement was observed between NRL3 and APN1 (κ = 0.878), while the lowest agreement was between NRL2 and CR (κ = 0.708). Hierarchical clustering based on the dissimilarity matrix (1 − κ) did not show a clear separation by professional group; instead, raters were interspersed across clusters in the clustered heatmap, suggesting broadly consistent interpretation across disciplines (Fig. 1).

Notes: NRL1/2/3: Neurologist 1/2/3, APN 1/2: Advanced Practice Nurse 1/2, CR: Clinical Researcher.

For the referral outcome, overall agreement across referral categories, estimated using Fleiss' κ, was statistically significant (κ = 0.668; Z = 69.6; $p < 0.001$). Category-specific values were as follows: Primary care (PC) κ = 0.661; Advanced Practice Nursing (APN) κ = 0.716; Outpatient specialist referral (OSR) κ = 0.309; Routine headache unit referral (RHU) κ = 0.321; and Preferential headache unit referral (PHU) κ = 0.820. The detailed results are presented in Supplementary Tables S3, S4 and Supplementary Fig. 1.

Classification performance

Of the 300 participants included in the validation sample, 277 (92.3%) had complete data and were included in the classification performance analysis (Fig. 2). MIGRAGILE classification was compared with the clinical reference diagnosis of disabling migraine (high-frequency episodic migraine or chronic migraine) established by headache-specialist neurologists according to ICHD-3 criteria.

At the optimal cut-off defined by the Youden index, MIGRAGILE showed a sensitivity of 93.7% (95% CI 91.7–95.3) and a specificity of 94.0% (95% CI 90.2–96.8) for classifying patients with disabling migraine. The positive predictive value was 89.0% (95% CI 84.9–93.0), and the negative predictive value was 96.6%

(95% CI 94.3–97.4). The likelihood ratios were LR + = 15.6 and LR− = 0.07. The area under the ROC curve was 0.91 (95% CI 0.87–0.95) (Fig. 3). The 2×2 contingency table is presented in Table 3.

Among the 277 complete cases, MIGRAGILE yielded 11 false positive and 6 false negative classifications. All false negative cases corresponded to patients with HFEM according to the clinical gold standard in the absence of objective biomarkers. Of these, four were classified by MIGRAGILE as other headache and two as non disabling migraine. All false positive cases were classified by MIGRAGILE as disabling migraine, whereas the specialist diagnosis was low frequency episodic migraine in five cases, tension type headache in three cases, new daily persistent headache in one case, greater occipital neuralgia in one case, and cluster headache in one case.

Table 3
Classification 2×2 table: MIGRAGILE vs. the clinical reference diagnosis (gold standard)

MIGRAGILE positive	**Disabling migraine (HFEM/CM)**	**Non-disabling / other headache**
	89 (true positive)	**11 (false positive)**
MIGRAGILE negative	6 (false negative)	171 (true negative)
Notes: Gold standard = ICHD-3 diagnosis by headache-specialist neurologists; n = 277 complete cases.		

Classification performance by sex and age

Classification performance was evaluated in subgroups defined by sex (men vs women) and age (< 40 vs ≥ 40 years). Estimates of sensitivity and specificity remained high across all subgroups (Table 4). Diagnostic odds ratios were statistically significant within each stratum ($p < 0.001$).

Table 4
Classification performance by sex and age subgroups.

Metric	Women	Men	< 40 years	≥ 40 years
n	238	39	137	140
TN	148	24	79	91
FP	9	3	4	8
FN	3	1	3	1
TP	78	11	51	40
Sensitivity	0.963	0.917	0.944	0.976
Specificity	0.943	0.889	0.952	0.919
PPV	0.897	0.786	0.927	0.833
NPV	0.980	0.960	0.963	0.989
OR (95% CI)	437.6 (112.6–1622.4)	88.0 (8.20–944.3)	335.8 (72.1–1562.7)	455.0 (55.1–3760.0)
p-value	< 0.001	< 0.001	< 0.001	< 0.001

Notes: TN = true negative; FP = false positive; FN = false negative; TP = true positive; PPV = positive predictive value; NPV = negative predictive value; OR = odds ratio.

Convergent validity with PROMs

Convergent validity was evaluated by examining the association between the clinical classification generated by MIGRAGILE and two validated measures of headache impact and migraine-specific quality of life: the Headache Impact Test-6 (HIT-6) and the Migraine-Specific Quality of Life Questionnaire, version 2.1 (MSQ v2.1).

HIT-6. The distribution of HIT-6 impact categories differed significantly across MIGRAGILE-defined groups ($\chi^2 = 115.99$; df = 6; $p < 0.001$). The effect size was Cramér's V = 0.46. Severe impact was reported by 65.5% of patients classified as disabling migraine, compared with 15.2% of those classified as non-disabling migraine and 15.2% of those with other headaches. No patients in the disabling migraine group were classified in the no/minimal impact category (Table 5).

MSQ v2.1 (range 14–84; higher scores indicate worse quality of life). MSQ scores differed significantly across MIGRAGILE groups (Kruskal–Wallis H = 45.62; df = 2; $p < 0.001$), with a large effect size ($\varepsilon^2 = 0.159$). Median (IQR) MSQ scores were highest in disabling migraine (61 [47–68]), followed by non-disabling migraine (41 [32–48.5]) and other headaches (29 [17–40.2]), indicating progressively worse quality of life in line with increasing clinical severity as defined by MIGRAGILE (Table 5).

Table 5
Convergent validity: PROMs according to MIGRAGILE classification. **MSQ v2.1**

Group	n	Mean ± SD	Median (IQR)	Range
Non-disabling migraine	157	41.98 ± 12.61	41 (32–48.5)	23–72
Disabling migraine	88	58.33 ± 13.71	61 (47–68)	23–84
Other headaches	32	29.3 ± 10.53	29 (17–40.2)	14–43
Notes: MSQ v2.1 (14–84; higher scores indicate worse quality of life); Global comparison: Kruskal–Wallis H = 45.62; $p < 0.001$; $\varepsilon^2 = 0.159$				

HIT-6 category distribution

Category	Non-disabling n (%)	Disabling n (%)	Other n (%)
No/minimal	81 (51.6)	0 (0.0)	17 (51.6)
Some impact	35 (22.1)	5 (5.8)	7 (22.1)
Substantial	17 (11.2)	25 (28.8)	3 (11.2)
Severe	24 (15.2)	58 (65.5)	5 (15.2)
Notes: Global comparison: $\chi^2(6) = 115.99$; $p < 0.001$; Cramér's V = 0.46			

Clinical pathways and therapeutic decisions

Among the 277 participants with complete data, 157 (56.7%) were classified as non-disabling migraine, 88 (31.8%) as disabling migraine (HFEM/CM), and 32 (11.6%) as other headaches. The distribution of care pathways differed significantly across MIGRAGILE groups ($\chi^2 = 234.8$; $p < 0.001$; Cramér's V = 0.65). Patients classified as disabling migraine were referred exclusively to the headache unit, whereas those classified as non-disabling migraine were managed predominantly through advanced practice nurse follow-up and primary care. In the other headaches group, referral to the headache unit was the predominant pathway. First-visit management decisions also differed significantly across groups ($p < 0.001$): non-specific oral preventive treatments and CGRP-targeted therapies were initiated mainly in patients classified as disabling migraine. In contrast, group therapeutic education was delivered exclusively to patients classified as non-disabling migraine ($p < 0.001$). Detailed results are presented in Table 6.

Table 6
Clinical pathways and therapeutic decisions according to MIGRAGILE classification (n = 277).

Variable	Non-disabling migraine (n = 157)	Disabling migraine (n = 88)	Other headaches (n = 32)	p-value
Care pathway				**< 0.001***
Headache unit – preferential	1 (0.6)	56 (63.6)	17 (53.1)	
Headache unit – ordinary	20 (12.7)	32 (36.4)	13 (40.6)	
APN follow-up	97 (61.8)	0 (0.0)	1 (3.1)	
Primary care follow-up |	39 (24.8)	0 (0.0)	1 (3.1)	
Treatment at first visit				**< 0.001****
Non-specific oral preventives	9 (5.7)	53 (60.2)	0 (0.0)	
CGRP-targeted treatments†	0 (0.0)	35 (39.8)	1 (3.1)	
Non-pharmacological intervention				**< 0.001***
Group therapeutic education	97 (61.8)	0 (0.0)	0 (0.0)	

Notes: Values are presented as n (%). *Chi-square test. **Fisher's exact test. †Includes monoclonal antibodies and onabotulinumtoxinA. APN: Advanced Practice Nurse. Cramér's V for care pathway = 0.65.

DISCUSSION

This study reports the development and validation of MIGRAGILE, a brief, self-administered digital questionnaire designed to identify, in primary care, patients likely to have disabling migraine—including high-frequency episodic migraine and chronic migraine—and to support structured referral decisions to specialised care. The findings show that MIGRAGILE has high classification performance for identifying patients likely to have disabling migraine, with strong sensitivity and specificity, good discriminative ability, and robust reliability indices, supporting its use as a patient-reported tool for clinical stratification in routine care.

Beyond classification performance, a key finding is that MIGRAGILE classification was associated with differentiated care pathways and therapeutic decisions consistent with patients' clinical complexity. Patients classified as having disabling migraine were predominantly referred to specialised services and more frequently initiated CGRP-targeted preventive therapies. These results suggest that MIGRAGILE may help guide clinical decisions from the first point of contact and support appropriate prioritisation of specialist resources, potentially reducing delays in referral and treatment initiation.

MIGRAGILE classification also showed a significant association with validated measures of migraine-related impact and quality of life. Patients classified as disabling migraine reported higher impact on the HIT-6 and worse MSQ scores, with large effect sizes. This consistency between MIGRAGILE clinical stratification and patient-reported disease burden provides additional evidence of convergent validity.

Within the landscape of existing migraine screening instruments, MIGRAGILE occupies a distinct position in terms of both purpose and conceptual design. Tools such as ID-Migraine™ and HARDSHIP [13, 15] have proven useful for population screening and epidemiological research, but they were not specifically developed to support clinical decision-making or structured care pathways. Other instruments, including VARS, MHQ-8 and MS-Q [14, 16, 17], address narrower dimensions of migraine, including visual aura features, sensory hypersensitivities associated with migraine attacks, and core screening criteria for migraine, respectively. As a result, their utility as comprehensive stratification tools in primary care is limited. By contrast, MIGRAGILE combines brevity with a targeted focus on identifying disabling migraine, in closer alignment with current diagnostic criteria and the organisational demands of healthcare systems.

This approach is particularly relevant in the current therapeutic context of migraine. The introduction of therapies targeting the calcitonin gene-related peptide (CGRP) pathway, including monoclonal antibodies, gepants and onabotulinumtoxinA, has expanded treatment options for patients with more severe phenotypes. In this setting, structured early identification of patients with greater clinical burden is essential to optimise treatment selection and reduce delays in access to effective preventive therapies.

From a measurement perspective, MIGRAGILE demonstrated excellent temporal stability for both categorical and quantitative items, as well as high inter-observer agreement in interpretation of outputs. This is especially relevant because MIGRAGILE was conceived for use in interprofessional care models, including pathways led by advanced practice nursing [43, 44]. The high agreement observed between neurologists and nurses suggests that the tool supports consistent interpretation of patient-reported clinical information and may facilitate coordination across levels of care.

High usability scores and short completion times support the feasibility of integrating MIGRAGILE into digital clinical workflows. In healthcare systems characterised by high demand and constrained resources, brief digital tools that combine patient-reported information with actionable clinical stratification may help improve organisational efficiency and prioritisation of specialist care [45, 46].

This study has limitations. Although validation was multicentre, it was conducted within a single healthcare system and geographic area in the city of Barcelona, which may limit generalisability to other contexts. In addition, MIGRAGILE was designed specifically to identify disabling migraine rather than to provide a comprehensive classification of all primary headache disorders. Future studies should evaluate external validity in other countries and care models, and examine longitudinal impact on clinical and organisational outcomes, including time to specialist referral, initiation of preventive treatment, and changes in patient-reported outcomes over time.

Overall, MIGRAGILE emerges as a brief, digital and psychometrically robust PROM that can support clinically meaningful stratification of patients in primary care. Its association with patient-reported burden, referral pathways and therapeutic decisions suggests potential utility in facilitating more appropriate referral to specialised care and more timely, targeted management in routine practice.

Electronic supplementary material

The link to the electronic supplementary material is provided below.

Abbreviations

APN
Advanced Practice Nurse
AUC
Area Under the Curve
CGRP
Calcitonin Gene–Related Peptide
CI
Confidence Interval
CM
Chronic Migraine
COSMIN
COnsensus–based Standards for the selection of health Measurement INstruments
CVI
Content Validity Index
EU
European Union
GCP
Good Clinical Practice
GDPR
General Data Protection Regulation
HARDSHIP
Headache–Attributed Restriction, Disability, Social Handicap and Impaired Participation (questionnaire)
HFEM
High–Frequency Episodic Migraine
HIT
6–Headache Impact Test–6
I
CVI–Item–level Content Validity Index
ICC
Intraclass Correlation Coefficient

ICHD
3–International Classification of Headache Disorders, 3rd edition
IQR
Interquartile Range
LR
Likelihood Ratio
LR+
Positive Likelihood Ratio
LR
–Negative Likelihood Ratio
MHD
Monthly Headache Days
MHQ
8–Migraine Hypersensitivity Questionnaire–8
MS
Q–Migraine Screen Questionnaire
MSQ
Migraine–Specific Quality of Life Questionnaire
NPV
Negative Predictive Value
OSR
Outpatient Specialist Referral
PC
Primary Care
PHU
Preferential Headache Unit
PPV
Positive Predictive Value
PRO
Patient–Reported Outcome
PROM
Patient–Reported Outcome Measure
RHU
Routine Headache Unit
ROC
Receiver Operating Characteristic
S
CVI/Ave–Scale–level Content Validity Index (Average method)
STARD
Standards for Reporting Diagnostic Accuracy Studies

SUS
System Usability Scale
VARS
Visual Aura Rating Scale

Declarations

Ethics approval and consent to participate

The study was conducted in accordance with Good Clinical Practice (GCP), the Declaration of Helsinki, and the EU General Data Protection Regulation (GDPR 2016/679). Ethical approval was obtained from the Clinical Research Ethics Committee of Hospital Clínic de Barcelona (reference number: HCB/2023/0606). All participants provided written informed consent prior to participation.

Competing Interests

TMG has received honoraria for participating in medical education activities from Eli Lilly, Organon, Novartis, Teva and Lundbeck. TMG has also received consulting fees from AbbVie and Lundbeck. In addition, TMG has received competitive public grants to fund nurse intensification contracts in the Strategic Plan for Health Research and Innovation 2022–2027 of the Government of Catalonia. VOB has received honoraria for participating in medical education activities from AbbVie, Eli Lilly, Organon, Novartis, Teva, Lundbeck and Dr. Reddy's. SFF has received honoraria for participating in medical education activities from Pfizer, Lundbeck and Organon, NFF has received honoraria for participating in medical education activities from Eli Lilly, Novartis, Teva, Lundbeck, Pfizer.

Funding

This research received no external funding.

Author Contribution

TMG and VOB conceived and designed the study and coordinated the study procedures. TMG was responsible for data collection, statistical analysis, and drafted the original manuscript. SFF, NFF, VOB, AGI and AS contributed to data interpretation, provided critical intellectual input, and participated in revising the manuscript for important scientific content. All authors contributed to editorial changes to the manuscript. All authors read and approved the final manuscript and agreed to its publication.

Acknowledgement

The authors would like to thank all participating patients and the healthcare professionals of the Barcelona Esquerra Integrated Health Area for their collaboration in this project.

Data Availability

The datasets supporting the findings of this study are held at the study institution and may be made available upon reasonable request and institutional approval; however, restrictions apply to their availability as they contain sensitive health information and were used under institutional authorisation for the current study and therefore are not publicly available. Nevertheless, data may be made available by the authors upon reasonable request and with approval from Hospital Clínic de Barcelona, in accordance with the EU General Data Protection Regulation (GDPR 2016/679) and institutional data access procedures.

References

1. Ashina M, Katsarava Z, Do TP, Buse DC, Pozo-Rosich P, Özge A et al (2021) Migraine: epidemiology and systems of care. Lancet 397:1485–1495. 10.1016/S0140-6736(20)32160-7
2. Dodick DW (2018) Migraine. Lancet 391(10127):1315–1330. 10.1016/S0140-6736(18)30478-1
3. García-Azorín D, Moya-Alarcón C, Armada B, Del Sánchez M (2024) Societal and economic burden of migraine in Spain: results from the 2020 National health and wellness survey. J Headache Pain 25:38. 10.1186/s10194-024-01740-3
4. Buse DC, Scher AI, Dodick DW, Reed ML, Fanning KM, Adams AM et al (2016) Impact of migraine on the family: perspectives of people with migraine and their spouse/domestic partner in the CaMEO study. Mayo Clin Proc 91(5):596–611. 10.1016/j.mayocp.2016.02.013
5. GBD 2016 Headache Collaborators (2018) Global, regional, and national burden of migraine and tension-type headache, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet Neurol 17(11):954–976. 10.1016/S1474-4422(18)30322-3
6. Evers S, Dell'Agnello G, Novick D, Gonderten HS, Panni T, Pascual J (2024) Acute treatment patterns, migraine burden, and healthcare resource use in people with migraine: results from the OVERCOME (EU) observational study. Pain Ther 13:589–607. 10.1007/s40122-024-00589-3
7. Pascual J, Núñez M, Panni T, Díaz-Cerezo S, Novick D, Ciudad A (2023) Burden and unmet needs in migraine patients: results from the OVERCOME (Spain) study. Pain Ther 12:1209–1220. 10.1007/s40122-023-00538-6
8. Sacco S, Amin FM, Ashina M, Bendtsen L, Deligianni CI, Gil-Gouveia R et al (2022) European Headache Federation guideline on the use of monoclonal antibodies targeting the calcitonin gene-related peptide pathway for migraine prevention—2022 update. J Headache Pain 23:67. 10.1186/s10194-022-01431-x
9. Pozo-Rosich P, Carmo M, Muñiz A, Armada B, Moya-Alarcón C, Pascual J (2024) Migraine treatment: quo vadis? Real-world data study (2015–2022) in Spain. BMC Neurol 24(1):107. 10.1186/s12883-

024-03600-8

10. Pascual J, Panni T, Dell'Agnello G, Gonderten S, Novick D, Evers S (2023) Preventive treatment patterns and treatment satisfaction in migraine: results of the OVERCOME (EU) study. J Headache Pain 24(1):88. 10.1186/s10194-023-01623-z
11. Porta-Etessam J, Santos-Lasaosa S, Rodríguez-Vico J, Núñez M, Ciudad A, Díaz-Cerezo S et al (2023) Evaluating the impact of a training program in shared decision-making for neurologists treating people with migraine. Neurol Ther 12(4):1319–1334. 10.1007/s40120-023-00495-4
12. Alpuente A, Gallardo VJ, Asskour L, Caronna E, Torres-Ferrus M, Pozo-Rosich P (2022) Salivary CGRP can monitor the different migraine phases: CGRP (in)dependent attacks. Cephalalgia 42(3):186–196. 10.1177/03331024211040467
13. Lipton RB, Dodick D, Sadovsky R, Kolodner K, Endicott J, Hettiarachchi J et al (2003) A self-administered screener for migraine in primary care: the ID Migraine validation study. Neurology 61(3):375–382. 10.1212/01.WNL.0000078940.53438.83
14. Láinez MJA, Domínguez M, Rejas J, Palacios G, Arriaza E, Garcia-Garcia M et al (2005) Development and validation of the Migraine Screen Questionnaire (MS-Q). Headache 45(10):1328–1338. 10.1111/j.1526-4610.2005.00265.x
15. Wang Y, Zhang Z, Wei J et al (2023) Validation of the HARDSHIP questionnaire for migraine diagnosis in nonclinical populations: a systematic review. Front Neurol 14:1231351. 10.3389/fneur.2023.1231351
16. Lee S, Kim H, Park JH et al (2023) Validation of the Visual Aura Rating Scale for migraine with aura diagnosis. Neurologia 38(4):253–260. 10.1016/j.nrl.2022.01.006
17. Gabrielli P, Rossi P, Di Fiore M et al (2025) Validation of an 8-item self-administered questionnaire for assessing migraine-related sensory hypersensitivities (MHQ-8). J Headache Pain 26(1):128. 10.1186/s10194-025-02067-3
18. Dikmen PY, Ozge A, Martelletti P (2023) The use of clinical scales and PROMs in headache and migraine disorders: summarizing their diffusion and operationalization. Heliyon 9(5):e16187. 10.1016/j.heliyon.2023.e16187
19. Rummel RJ (1970) Applied factor analysis. Northwestern University, Evanston
20. Gorsuch RL (1997) Exploratory factor analysis: its role in item analysis. J Pers Assess 68:532–560. 10.1207/s15327752jpa6803_5
21. Ten Berge JMF (1995) J. C. Nunnally and I. H. Bernstein. Psychometric Theory. New York: McGraw-Hill, 1994, xxiv + 752 pages. Psychometrika 60(2):313–315. 10.1007/BF02301419
22. Rovinelli RJ, Hambleton RK (1977) On the use of content specialists in the assessment of criterion-referenced test item validity. Dutch J Educ Res 2:49–60
23. Castilla D, Jaen I, Suso-Ribera C, Garcia-Soriano G, Zaragoza I, Breton-Lopez J et al (2024) Psychometric properties of the Spanish full and short forms of the System Usability Scale (SUS): detecting the effect of negatively worded items. Int J Hum Comput Interact 40(15):4145–4151. 10.1080/10447318.2023.2209840

24. Landis JR, Koch GG (1977) The measurement of observer agreement for categorical data. Biometrics 33(1):159–174
25. Koo TK, Li MY (2016) A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med 15(2):155–163. 10.1016/j.jcm.2016.02.012
26. Cohen J (1968) Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. Psychol Bull 70(4):213–220. 10.1037/h0026256
27. Sim J, Wright CC (2005) The kappa statistic in reliability studies: use, interpretation, and sample size requirements. Phys Ther 85(3):257–268
28. McHugh ML (2012) Interrater reliability: the kappa statistic. Biochem Med (Zagreb) 22(3):276–282
29. Fleiss JL (1971) Measuring nominal scale agreement among many raters. Psychol Bull 76(5):378–382. 10.1037/h0031619
30. McKenzie KA, Mahnken JD (2023) Simulating and estimating agreement in the presence of multiple raters and covariates. Stat Med 42(11):1687–1698. 10.1002/sim.9694
31. Zhou XH, Obuchowski NA, McClish DK (2011) Statistical methods in diagnostic medicine, 2nd edn. Wiley, Hoboken
32. Pepe MS (2003) The statistical evaluation of medical tests for classification and prediction. Oxford University Press, Oxford. 10.1093/oso/9780198509844.001.0001
33. Youden WJ (1950) Index for rating diagnostic tests. Cancer 3(1):32–35. 10.1002/1097-0142(1950)3:1%3C32:AID-CNCR2820030106%3E3.0.CO;2-3
34. Newcombe RG (1998) Two-sided confidence intervals for the single proportion: comparison of seven methods. Stat Med 17(8):857–872. 10.1002/(SICI)1097-0258(19980430)17:8%3C857:AID-SIM777%3E3.0.CO;2-E
35. Kosinski M, Bayliss MS, Bjorner JB, Ware JE Jr, Garber WH, Batenhorst A et al (2003) A six-item short-form survey for measuring headache impact: the HIT-6. Qual Life Res 12(8):963–974. 10.1023/A:1026119331193
36. Bagley CL, Rendas-Baum R, Maglinte GA, Yang M, Varon SF, Lee J, Kosinski M (2012) Validating Migraine-Specific Quality of Life Questionnaire v2.1 in episodic and chronic migraine. Headache 52(3):409–421. 10.1111/j.1526-4610.2011.01997.x
37. Cramér H (1946) Mathematical methods of statistics. Princeton University Press, Princeton, NJ
38. Cohen J (1988) Statistical power analysis for the behavioral sciences, 2nd edn. Lawrence Erlbaum, Hillsdale, NJ
39. Tomczak M, Tomczak E (2014) The need to report effect size estimates revisited: an overview of some recommended measures of effect size. Trends Sport Sci 21(1):19–25
40. Fritz CO, Morris PE, Richler JJ (2012) Effect size estimates: current use, calculations, and interpretation. J Exp Psychol Gen 141(1):2–18. 10.1037/a0024338
41. Shapiro SS, Wilk MB (1965) An analysis of variance test for normality (complete samples). Biometrika 52(3–4):591–611. 10.1093/biomet/52.3-4.591

42. Gagnier JJ, Lai J, Mokkink LB, Terwee CB (2021) COSMIN reporting guideline for studies on measurement properties of patient-reported outcome measures. Qual Life Res 30(8):2197–2218. 10.1007/s11136-021-02822-4
43. Rasmussen AV, Jensen RH, Gantenbein A et al (2024) Consensus recommendations on the role of nurses in headache care: a European e-Delphi study. Cephalalgia 44(5):3331024241252161. 10.1177/03331024241252161
44. Bhola R, Ertem DH (2022) The role and impact of the headache nurse specialist. Agri 34(1):75–76. 10.14744/agri.2020.67365
45. Topol EJ (2019) High-performance medicine: the convergence of human and artificial intelligence. Nat Med 25(1):44–56. 10.1038/s41591-018-0300-7
46. Bashshur RL, Howell JD, Krupinski EA, Harms KM, Bashshur N, Doarn CR (2016) The empirical foundations of telemedicine interventions in primary care. Telemed J E Health 22(5):342–375. 10.1089/tmj.2016.0045

Figures

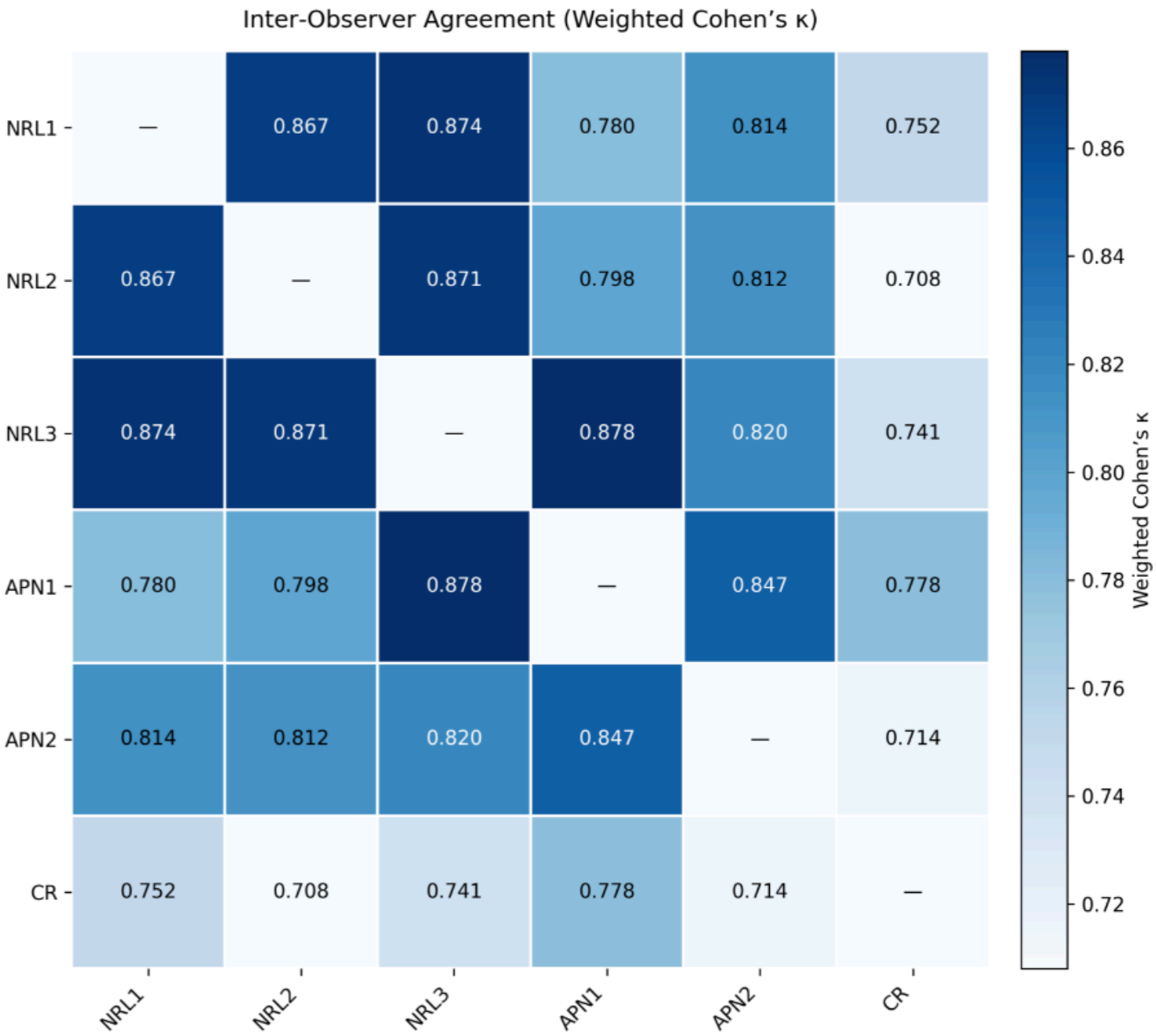


Figure 1

Clustered heatmap of inter-observer agreement (weighted Cohen's κ).
Dendrogram representing hierarchical clustering of observers based on the dissimilarity matrix (1 − κ) derived from pairwise weighted Cohen's κ coefficients for the reason-for-consultation output. Shorter branch lengths indicate higher agreement between raters.

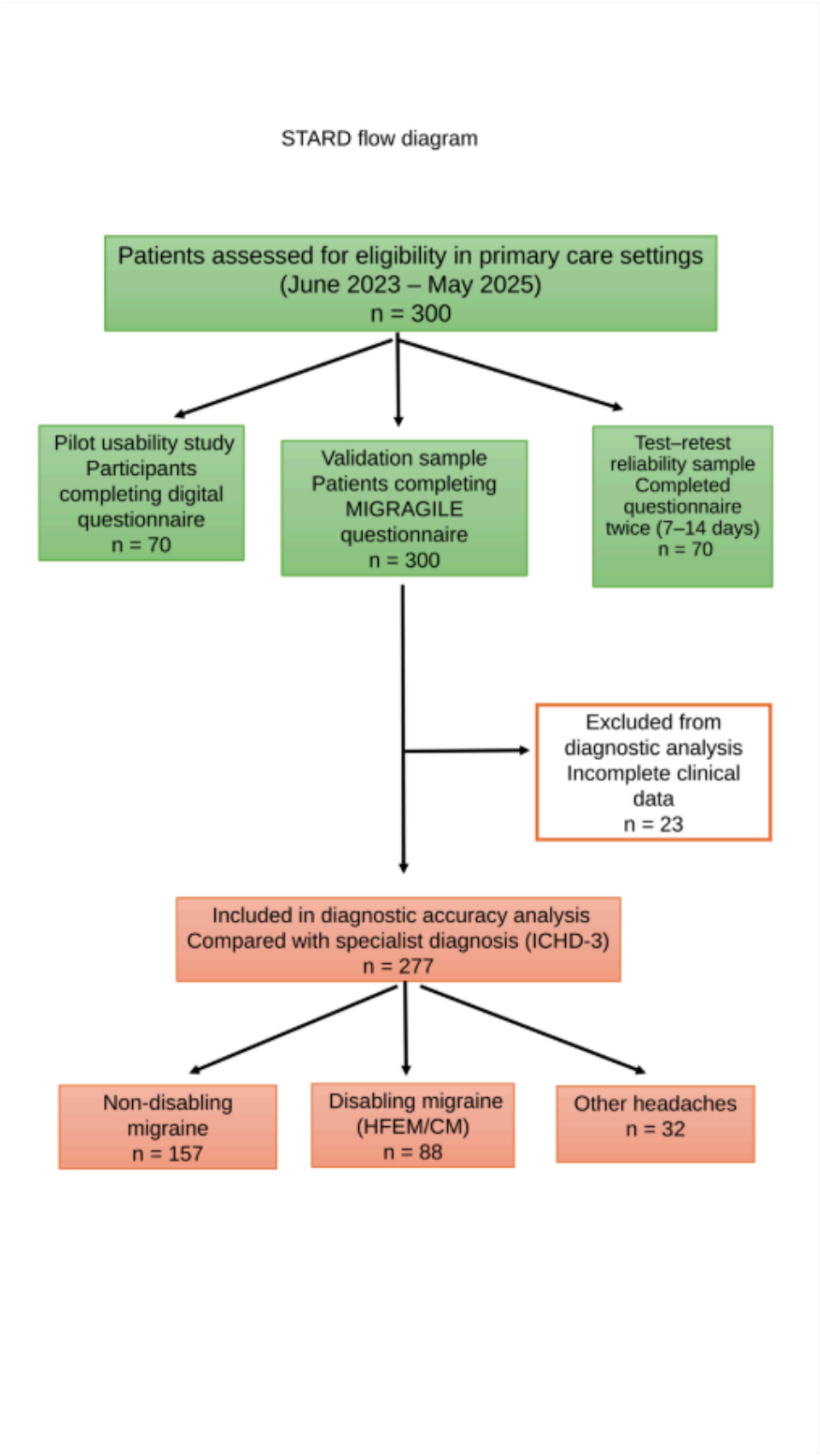


Figure 2

STARD flow diagram.

Flow diagram illustrating patient recruitment, index test administration (MIGRAGILE), reference standard assessment (ICHD-3 specialist diagnosis), and inclusion in the classification performance analysis. Pilot usability (n=70) and test–retest reliability subsamples (n=70) are shown as nested samples from the index-test cohort (N=300).

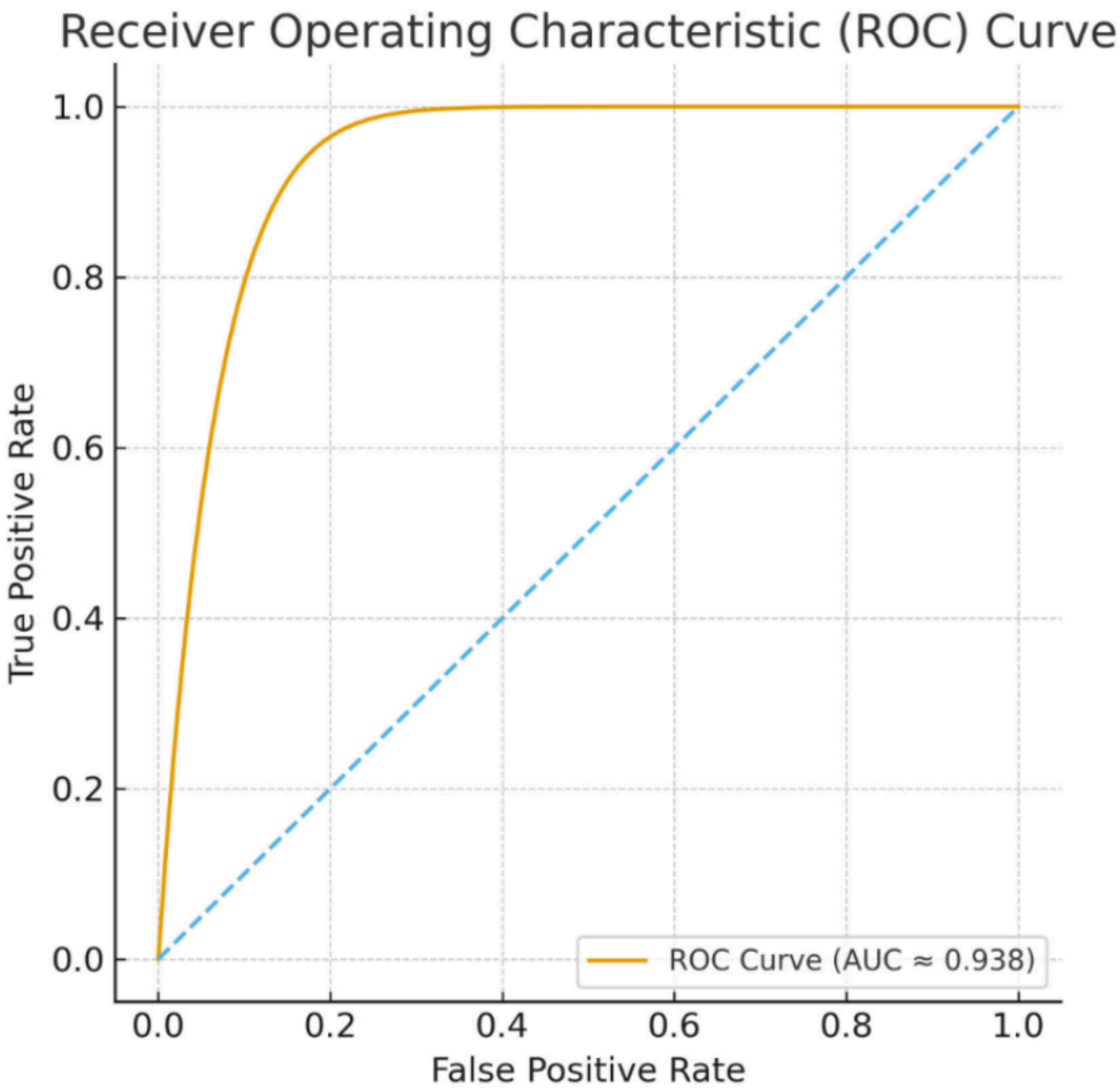


Figure 3

Receiver operating characteristic (ROC) curve.
Receiver operating characteristic (ROC) curve showing the ability of MIGRAGILE to discriminate disabling migraine (HFEM/CM) from non-disabling migraine/other headaches. The AUC and 95% CI are reported in the results.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- MIGRAGILESupplementaryMaterial.docx