Research Submission

Evaluating the Psychometric Properties of the Migraine Functional Impact Questionnaire (MFIQ)

Ariane K. Kawata, PhD; Asha Hareendran, PhD; Shannon Shaffer, MS; Sally Mannix, BA; Andrew Thach, PhD; Pooja Desai, PhD; Daniel D. Mikol, MD, PhD; Brian Ortmeier, PharmD, PhD; Martha Bayliss, MSc; Dawn C. Buse, PhD

Background.—Migraine is a chronic neurologic disease that can be associated with significant migraine-related impact, disability, and burden. Patient-reported outcome measures (PRO) are included in clinical trials of migraine interventions to capture treatment effects from a patient perspective. Clinical and regulatory guidelines also encourage use of PROs in trials. The Migraine Functional Impact Questionnaire (MFIQ) is a novel PRO measure, assessing the impact of migraine on Physical Function (PF), Usual Activities (UA), Social Function (SF), and Emotional Function (EF), in the past 7 days. Scientific methods recommended to meet the requirements of the U.S. Food and Drug Administration were followed, to ensure that the MFIQ content included outcomes that were relevant to adults with migraine and were clinically relevant, specifically to evaluate preventive treatments for migraine.

Objective.—The objective of this study was to conduct item analyses informing item reduction and scoring, and to evaluate the psychometric properties of the MFIQ.

Methods.—In a prospective, observational study, adults with migraine completed the MFIQ as well as additional clinical and PRO instruments, including the Headache Impact Test (HIT-6^{TM}), Patient-Reported Outcomes Measurement Information System Physical Function Short Form 10a (PROMIS-PF), Migraine-Specific Quality-of-Life Questionnaire (MSQ), and Patient Global Rating of Change (PGIC). Item-level evaluation, item response theory (IRT), and factor analysis were used to select final MFIQ items, identify domains, and inform scoring. Psychometric properties of the MFIQ were evaluated to assess reliability (internal consistency and test–retest), validity (construct and known-groups), and responsiveness.

Results.—The study enrolled 569 adults with migraine. Subjects had an average age of 39.9 years (SD 12.0), 87.2% were female, and 80.8% were white. Five items were dropped from the draft version based on results of item-level analyses reviewed in the context of previous qualitative research to produce the final 26-item MFIQ (v.2). Four domain scores (PF, UA, SF, and EF) and a global item score for impact on UA were identified. Higher scores on a 0-100 scale represent greater impact. All scores exhibited high internal consistency ($\alpha \geq 0.89$) and moderate test–retest reliability among stable subjects (ICCs ≥ 0.47). Construct validity was demonstrated by significant correlations (all $P < .0001$) between MFIQ domain scores, related PRO scores, and the frequency of migraine days and headache days. All domain scores differentiated between subgroups ("known groups") defined based on established levels of clinical severity: number of monthly migraine and headache days, migraine interference levels and scores on other PRO instruments ($P < .05$). Improvements in MFIQ scores corresponded with clinical improvement (percent reduction in monthly migraine days), improvement in migraine interference with daily activities, and related improvements in PRO scores ($P < .05$), demonstrating that the MFIQ was responsive to changes in migraine impact.

From Evidera, Bethesda, MD, USA (A.K. Kawata, S. Shaffer, and S. Mannix); Evidera, London, UK (A. Hareendran); Amgen Inc., Thousand Oaks, CA, USA (A. Thach, P. Desai, D.D. Mikol, and B. Ortmeier); Optum, Lincoln, RI, USA (M. Bayliss); Department of Neurology, Albert Einstein College of Medicine, Bronx, NY, USA (D.C. Buse).

Address all correspondence to A.K. Kawata, Evidera, 7101 Wisconsin Avenue, Suite 1400, Bethesda, MD 20814, USA, email: ariane.kawata@evidera.com

Conclusions.—The MFIQ is a reliable and valid measure that can be used to collect data about migraine impact. The MFIQ is being used to evaluate outcomes of migraine interventions in clinical trials and observational studies. It could potentially also be used in clinical practice both for initial and ongoing assessments for monitoring outcomes and to enhance communication between patients and healthcare professionals for the management of migraine.

Key words: migraine, reliability, validity, responsiveness, functioning, MFIQ

Abbreviations: CFA confirmatory factor analysis, CM chronic migraine, EF Emotional Function, EFA exploratory factor analysis, EM episodic migraine, GRM graded response model, HIT-6™ Headache Impact Test, ICC intraclass correlation coefficient, IRT item response theory, MFIQ Migraine Functional Impact Questionnaire, MSQ Migraine-Specific Quality-of-Life Questionnaire, PGIC Patient Global Impression of Change, PRO patient-reported outcome, PROMIS-PF Patient Reported Outcomes Measurement Information System Physical Function Short Form 10a, RF-P Role Function-Preventive, RF-R Role Function-Restrictive, RMSEA root mean square error of approximation, SRMR standardized root mean residual, WLSMV weighted least squares mean and variance adjusted

INTRODUCTION

Migraine is a chronic neurologic disease that is known to cause debilitating and disruptive physical symptoms that often impact social, emotional, academic, occupational, and personal aspects of life.[1-5] U.S. Headache Consortium Guidelines for the Pharmacological Management of Migraine Headache in the Primary Care Setting recommend that treatments for migraine should improve function and reduce disability, in addition to reducing the frequency, intensity, and duration of attacks and improving responsiveness of attacks to treatment.[6] Guidelines for the management and evaluation of migraine treatments also emphasize the importance of collecting data on the impact of migraine on functioning (ie, reducing migraine-related disability, burden, and impact), and suggest using patient-reported outcome (PRO) instruments to support clinical endpoints in migraine clinical trials.[7,8] Guidelines recommend that PRO instruments should include concepts relevant to patients (using evidence based on patient input) and have robust measurement properties in the target population.[9]

Several PRO instruments exist that measure migraine-related disability, burden, and impact, such as the Headache Impact Test (HIT-6™),[10] Migraine Disability Assessment (MIDAS),[11] and Migraine-Specific Quality-of-Life Questionnaire (MSQ).[12] However, a review showed that these instruments had limited coverage of some of the multifaceted impacts of migraine on functioning, that are relevant to adults, particularly the key facets of physical functioning: **acts** (things that an individual can do independent of context or purpose) and **tasks** (things people do in daily life in a specific context, with purpose).[13] Existing instruments also have gaps in FDA-recommended development steps required by regulatory guidelines issued in 2009 that highlight the importance of content validity and patient input.[1,9,14] The existing instruments also utilize longer recall periods that capture the patient's overall assessment of their condition but not recent episodes.[15] For example, the HIT-6™, MIDAS, and MSQ have recall periods of 4 weeks or longer. In contrast, a daily diary such as the Migraine Physical Function Impact Diary (MPFID),[16,17] which captures the daily impact of migraine, can ideally track the day-to-day fluctuations

that can occur during the various phases of a migraine episode. A recall period longer than 7 days may potentially limit the accuracy of recall about these day-to-day variations and impacts of migraine and related symptoms.

The Migraine Functional Impact Questionnaire (MFIQ) was developed to address the gaps in existing PROs, specifically by providing a method to capture the comprehensive impact of migraine in the *past 7 days* on Physical Function (PF); Usual Activities (UA); Social Function (SF), and Emotional Function (EF). A 7-day recall period was selected to capture the variability of migraine symptoms and impacts, while being less burdensome than a daily assessment and reducing potential recall bias associated with longer recall intervals. The initial MFIQ items were generated based on concept elicitation interviews with patients, and content validity was further confirmed via cognitive interviews.[1,14] Clinical experts in migraine also provided input in the development and finalization of the MFIQ. Items use a 5-point response scale. For items relating to concepts/activity that may not be relevant to the subject during a particular week (eg, "ability to take care of your family)," a "does not apply" response option can be used.

The MFIQ was developed for use in clinical trials and other research settings. It has recently been used in trials evaluating an investigational migraine preventive therapy[18] and a prospective observational survey of migraine preventive treatment.[19,20] The MFIQ has also been translated into 20 languages following best practice recommendations for linguistic validation of PRO instruments.[14] The MFIQ has also been noted in the recently published "Guidelines of the International Headache Society for controlled trials of preventive treatment of chronic migraine in adults" as a PRO tool that can be used to support secondary endpoints in clinical trials of preventive treatments for migraine.[21]

Strong measurement properties of a PRO instrument are crucial for the interpretability and generalizability of the constructs being measured.[22] The primary objective of this study was to finalize the content of the MFIQ and conduct a psychometric evaluation of measurement properties using data from an observational study.

METHODS

Study Design and Population.—A prospective, observational study was conducted (in 2014-2016) with adults 18 to 65 years of age diagnosed with migraine, who could speak and read English. Eligibility criteria aimed to mirror criteria typically used for a migraine clinical trial of preventive therapies, to capture a similar patient population. The study enrolled adults with episodic migraine (≥4 and ≤14 migraine headache days per month, in each of 3 months prior to screening) and chronic migraine (≥15 headache days per month, of which ≥8 were migraine days, in each of the 3 months prior to screening), consistent with ICHD-III 2013 criteria for EM and CM. The study excluded individuals participating in a clinical trial within 45 days of screening; older than 50 at time of migraine onset; people with cluster headache or hemiplegic migraine headache, or with comorbid conditions known to cause substantial pain (fibromyalgia, chronic pelvic pain syndrome); EM with migraine attacks lasting over 72 hours in duration in the 3 months prior to screening; CM with continuous pain; opioid use for greater than 6 days per month during the 3 months prior to screening; individuals unable to differentiate migraine from other headache; and anyone with a cognitive impairment that would preclude participation (in the opinion of the investigator). Diagnoses were confirmed by clinicians (nurses, doctors, and other healthcare professionals) or via chart review and patient self-report of number of migraine days during screening.

The study included 2 cohorts. Cohort 1 (followed for 4 weeks) was used to inform item reduction, scoring, and cross-sectional properties (reliability and validity) and included subjects who were receiving usual standard of care migraine treatment, either taking a preventive migraine medication or no preventive treatment (acute treatments were permitted). Cohort 2 (followed for 16 weeks) was used to examine responsiveness to change included subjects starting a new preventive medication and/or increasing the dose of a current preventive medication.

Note: Data from this observational study were also used to explore the properties of a daily assessment of the impact of migraine, specifically on physical function using the MPFID. Information about the MPFID is reported in Hareendran et al 2017[16] and Kawata, Hsieh, et al 2017.[17]

Subjects were recruited via 35 clinical sites in geographically diverse areas of the United States, including general practice facilities and neurology/headache centers. The sample size for Cohort 1 was determined based on a minimum requirement of at least 5-10 subjects per item or 100-200 subjects[23,24] for item evaluation and factor analyses. The sample size for Cohort 2 was estimated for evaluating responsiveness based on the results of 2 pivotal trials of topiramate[25] which evaluated the change in average area under the curve (AUC) of the MSQ. A sample size of at least 100 subjects was required for Cohort 2, to include an adequate number of subjects for a preliminary evaluation of longitudinal properties to discriminate responder and nonresponder groups. A type I error rate of $P < .05$ was assumed, and a 2-sided 2-sample *t*-test was used for the power calculation.

All study procedures were approved by a central Institutional Review Board (Ethical & Independent Review Services, Corte Madera, CA, 14020-01), and subjects provided written informed consent. Subjects attended 2 clinic visits: Cohort 1 at Day 1 (study start) and Week 4, and Cohort 2 at Day 1 (study start) and Week 16, and responded to questions on an electronic device each day between study visits. Subjects completed a headache diary on the electronic device that collected data about headache and associated symptoms, medications taken for acute migraine pain and symptoms, and the level of migraine interference with daily activities (average score, categorized as none [0], mild [1-3], moderate [4-6], severe [7-10]). PRO instruments were also completed on the electronic device throughout the study at specific time points (Table 1). These instruments included MFIQ (v.1); HIT-6[TM;10,26] Patient-Reported Outcomes Measurement Information System Physical Function Short Form 10a (PROMIS-PF);[27] MSQ (v2.1);[12,28] and a Patient Global Rating of Change (PGIC) question. Table 1 illustrates the schedule of instrument completion in each cohort and scores generated from these instruments.

Item analyses were conducted on the 31-item draft MFIQ (v.1), and psychometric analyses were conducted with the subset of items retained in the final MFIQ (v.2; 26 items). Version 1.0 and the final version of the MFIQ representing impacts of migraine in 4 areas of functioning is described in Table 2.

Statistical Analyses.—Statistical analyses were conducted in 3 stages and in accordance with best practices for instrument development/psychometric evaluation[23] and recommendations outlined in the FDA PRO Guidance document.[9]

In Stage 1, an interim sample (n = 259) was used to evaluate the performance of items in MFIQ v.1 at study start (Day 1) to confirm content validity, inform item reduction, identify domain structure, and develop a scoring algorithm utilizing data collected at study start from Cohorts 1 and 2. Following completion of item reduction and scoring in Stage 1, the measurement properties of MFIQ v.2 were evaluated using the full study sample. In Stage 2 (n = 553), the cross-sectional measurement properties of the domains identified in Stage 1 were assessed, based on data collected at the end of Week 4 from Cohorts 1 and 2. In Stage 3 (n = 267), responsiveness of MFIQ v.2 scores were assessed with the Cohort 2 sample using analysis of covariance (ANCOVA) models to compare score change from Weeks 4 or 8 to Week 16. Time points used to evaluate change over time in Cohort 2 were adjusted based on the assessment schedule (see Table 1).

Analyses were conducted using all available data and missing clinical and PRO data were not imputed. All statistical tests (unless otherwise noted) used a significance level of $P < .05$ (2-sided) and did not adjust for multiplicity. All statistical analyses were performed using SAS 9.4 (SAS Institute Inc., Cary, NC), with the exceptions of factor analyses conducted using Mplus 7.11[29] and item response theory (IRT) analyses performed using IRTPRO 2.1.[30]

Stage 1: Item Analysis, Item Reduction, and Scoring.—Item-level descriptive statistics were generated for the MFIQ v.1. *A priori* criteria for flagging items for potential item reduction included: large floor or ceiling effects (>30% at minimum or maximum response option); high ($r > 0.80$) or low ($r < 0.20$) item-to-item correlations; low factor loadings (<0.40) or items that loaded on more than 1 factor (>0.40); and misfit within the domain in IRT analysis (item chi-square $P < .01$[30] or based on Bonferroni-adjusted criteria).

Item selection and scoring were informed by exploratory factor analysis (EFA) and IRT analysis using the graded response model. IRT was used

Table 1.—Schedule of Assessments

| PRO Instrument | | Study Schedule | | | | | |
| | | 1 = Cohort 1; 2 = Cohort 2 | | | | | |
Name and Recall Period	Range of Scores and Interpretation	Day 1/Study Start	Week 2	Week 4	Week 8	Week 12	Week 16
MFIQ Past 7 days	Physical Function Usual Activities Social Function Emotional Function 0-100; higher scores = greater impact	1,2		1,2	2	2	2
HIT-6™ Past 4 weeks and current	Severity of headache impact 36-78; higher scores = greater impact Minimal impact <50, mild impact 50-55, moderate impact 56-59, Severe impact >59	1,2		1	2		2
MSQ Past 4 weeks	Role Function-Restrictive Role Function-Preventive Emotional Function 0-100; higher scores = better quality of life	1,2		1	2		2
PROMIS-PF Current	Ability to carry out physical activities 0-30; higher scores = better able to carry out activities	1,2		1	2		2
PGIC Since start of study (day 1)	Overall assessment of change in migraine 7-point scale; very much worse to very much better		1,2	1,2	2	2	2

HIT-6™ = 6-item Headache Impact Test; MFIQ = Migraine Functional Impact Questionnaire; MSQ = Migraine-Specific Quality-of-Life Questionnaire; PGIC = Patient Global Rating of Change in Migraine; PRO = Patient reported outcome; PROMIS-PF = Patient-Reported Monitoring Information System-Physical Function 10-item Short Form.

Table 2.—Content of the MFIQ Version 1.0 and Version 2.0

Domain of Functioning	Number of Items	Concepts Covered
Impact on Physical Function	6	**Frequency of impact on:** ability to move head, ability to move body, *ability to move inside home*,† ability to do activities needing physical effort, needing to rest or lie down, feeling too tired to do things
Impact on Usual Activities	13	**Level of difficulty:** to get ready for the day, do usual activities, *activities with others*,† do usual household chores, do chores outside of the home, do activities that require concentration, *do activities that require thinking clearly*,† school or work, take care of family, impact on activities in extreme sensations (sound, smell, light) **Frequency of difficulty:** completing personal grooming activities, affecting daily routine or schedule, having to change plans
Impact on Social Function	7	**Frequency of impact on:** social interactions, being around other people, participating in social activities, relationships, *hobbies*,† leisure activities, *talking with family, friends, coworkers*†
Impact on Emotional Function	5	**Frequency of feeling:** frustrated because of migraine, feeling worried because of migraine, like a burden on others because of migraine, feeling lack of control of life because of migraine, disappointed because of a migraine

†Items in *italics* were deleted for Version 2.0 based on item analyses.

as a secondary method to assess item performance and unidimensionality of the MFIQ domains in conjunction with item-level evaluation and factor analysis. Confirmatory factor analysis (CFA) was used to confirm the factor structure derived using EFA. EFA and CFA models were conducted using weighted least squares mean and variance adjusted (WLSMV) estimation and goodness of fit was assessed based on comparative fit index (CFI; ≥0.90),[31] root mean square error of approximation (RMSEA; <0.08 with narrow 90% confidence interval),[32] and standardized root mean residual (SRMR; <0.10) for EFA or weighted root mean square residual (WRMR; close to or ≤1.0)[33] for CFA.

Stage 2: Cross-Sectional Properties: Reliability and Validity.—The following measurement properties of the MFIQ v.2 were evaluated.

Internal consistency reliability (the extent to which individual items in a scale measure a common underlying concept) was evaluated at study start (Day 1) and Week 4 using Cronbach's alpha >0.70 as indicative of acceptable reliability.[34] *Test–retest reliability* (ability to generate reproducible results) of MFIQ scores was assessed among subjects who had stable status (based on PGIC "no change in migraine" at Week 4) from Day 1 to Week 4 using intraclass correlation coefficients (ICC). An ICC >0.70 indicates good test–retest

reliability, between 0.4 and 0.7 indicates moderate reliability, and <0.4 indicates low reliability.[23]

Construct validity was evaluated with data from Week 4 using Spearman rank-order correlations. Correlations between MFIQ scores and other indicators of conceptually related constructs based on clinical and PRO assessments that would be expected to show associations were examined to support convergent validity. For example, moderate correlations were expected between MFIQ scores and frequency of migraine days and headache days; moderate to large correlations were expected between MFIQ PF and PROMIS-PF scores, between MFIQ UA and SF scores with MSQ Role Function-Restrictive (RF-R) and Role Function-Preventive (RF-P) scores, and between MFIQ EF and MSQ EF scores. Strength of a correlation was evaluated based on standards proposed by Hinkle and colleagues: small <0.3; moderate 0.3-0.6; large >0.6.[35]

Known-groups validity (the extent to which scores show expected differences among groups of subjects differing by a relevant clinical marker[36] was assessed using data from Week 4 based on groups known to be clinically different based on headache variables and other PRO scores: number of weekly migraine days and headache days (categorized as 0 days, 1-2 days, ≥3 days), level of migraine interference with daily

Table 3.—Sociodemographic and Clinical Characteristics

		Subsets	
	Study Sample for Stage 2 Analyses (Cohorts 1 and 2)	Interim Sample for Stage 1 Analyses (Cohorts 1 and 2)	Sample for Stage 3 Analyses (Cohort 2)
N at study start	569	264	308
Age (years)			
N, Mean (SD), Min–Max	562, 39.9 (12.0), 18-65	262, 40.6 (11.3), 18-64	304, 39.6 (12.3), 18-65
Female n, %	496 (87.2%)	229 (86.7%)	270 (87.7%)
Hispanic or Latino n, %	58 (10.2%)	33 (12.9%)	24 (7.8%)
Race† n, %			
Asian	7 (1.2%)	3 (1.1%)	4 (1.3%)
Black or African American	80 (14.1%)	43 (16.3%)	39 (12.7%)
White	460 (80.8%)	203 (75.9%)	265 (86.0%)
Other	33 (5.8%)	22 (8.3%)	15 (4.9%)
Employment status† n, %			
Employed full- or part-time	399 (70.1%)	187 (70.8%)	218 (70.8%)
Student	64 (11.2%)	26 (9.8%)	33 (10.7%)
Homemaker	53 (9.3%)	23 (8.7%)	32 (10.4%)
Unemployed/retired/disabled	97 (17.0%)	49 (18.6%)	50 (16.2%)
Other	3 (0.5%)	2 (0.8%)	1 (0.3%)
Education level n, %			
Secondary/high school or less	80 (14.0%)	38 (14.4%)	41 (13.3%)
Some college	193 (33.9%)	82 (31.1%)	112 (36.4%)
College or postgraduate degree	278 (48.8%)	136 (51.5%)	150 (48.7%)
Other	9 (1.6%)	6 (2.3%)	5 (1.6%)
Clinical characteristics			
Migraine classification n, %			
Episodic migraine	323 (56.8%)	181 (68.6%)	181 (58.8%)
Chronic migraine	246 (43.2%)	83 (31.4%)	127 (41.2%)
Migraine without aura‡ n, % yes	367 (64.5%)	166 (62.9%)	214 (69.5%)
Migraine with aura‡ n, % yes	238 (41.8%)	118 (44.7%)	117 (38.0%)
Menstrual migraine‡ n, %	162 (28.5%)	77 (29.2%)	82 (26.6%)

†Check all that apply; percentages may sum to more than 100%.
‡Clinician-reported; diagnoses of migraine without aura and migraine with aura are not mutually exclusive.
Min–Max = minimum–maximum; SD = standard deviation.

activities, HIT-6 severity of headache impact score, and MSQ domains (median scores). Mean MFIQ scores and standard deviations (SDs) were calculated for each severity/disease status category and ANCOVA was used to compare mean differences between groups, controlling for age and sex.

Stage 3: Responsiveness.—In Stage 3, the ability of the MFIQ v.2 to detect change (*responsiveness*) from Week 4 or Week 8 to Week 16 (n = 267) was assessed. ANCOVA compared MFIQ score changes over time (controlling for age and sex), and baseline MFIQ scores between groups differing by responder status. A responder was defined as a subject showing change on clinical variables (≥50% reduction in monthly migraine days or monthly headache days) and PRO anchor variables (≥50% reduction in migraine interference with daily activities; improved on the PGIC; change of at least 5 points on HIT-6; change of at least 5 points on MSQ RF-R and RF-P domains, and change of at least 8 points on MSQ EF domain).

RESULTS

Patient Characteristics.—The full study sample used in Stage 2 to evaluate psychometric properties of the MFIQ included 569 subjects: 323 (56.8%) with EM and

Table 4.—Patient-Reported Assessments of Headache Impact and Functioning

| Measure | Study Sample for Stage 2 Analyses (Cohorts 1 and 2) | Subsets | |
		Interim Sample for Stage 1 Analyses (Cohorts 1 and 2)	Sample for Stage 3 Analyses (Cohort 2)
N	569	262	308
HIT-6 Score[†], Mean (SD)	63.6 (5.9)	63.6 (5.7)	63.5 (6.2)
HIT-6 Score Categories, n (%)			
Minimal Impact (<50)	6 (1.1%)	1 (0.4%)	6 (1.9%)
Mild Impact (50-55)	39 (6.9%)	21 (8.0%)	17 (5.5%)
Moderate Impact (56-59)	54 (9.5%)	22 (8.4%)	35 (11.4%)
Severe Impact (>59)	449 (78.9%)	210 (80.2%)	244 (79.2%)
Missing	21 (3.7%)	8 (3.1%)	6 (1.9%)
MSQ Domain Scores[‡], Mean (SD)			
Role Function-Restrictive	47.6 (20.9)	47.8 (21.1)	47.1 (20.8)
Role Function-Preventive	63.4 (24.1)	64.6 (24.2)	62.4 (23.8)
Emotional Function	54.2 (29.0)	56.7 (28.1)	51.6 (28.9)
PROMIS-PF T-score[§], Mean (SD)	49.2 (8.5)	49.7 (8.6)	49.0 (8.6)

[†]HIT-6 scores range from 36 to 78; higher scores indicate greater impact of headache.
[‡]MSQ domain scores range from 0 to 100; higher scores indicate better migraine-specific quality of life.
[§]PROMIS-PF Short-Form T-scores are standardized with a mean of 50 and SD of 10, where higher scores indicate better physical functioning.
HIT-6 = 6-item Headache Impact Test; MSQ = Migraine-Specific Quality-of-Life Questionnaire; PROMIS-PF = Patient-Reported Outcomes Measurement Information System-Physical Function; SD = standard deviation.

246 (43.2%) with CM. The mean age was 39.9 years (SD = 12.0; range: 18-65), and 87.2% (n = 496) were female, 80.8% (n = 460) were white, and most (70.1%; n = 399) were employed full- or part-time. Clinicians categorized 64.5% (n = 367) of subjects as having migraine without aura. The most common current migraine preventive medications were topiramate (34.3%; n = 195) and tricyclic antidepressants (eg, amitriptyline, nortriptyline, protriptyline) (17.6%; n = 100). Approximately one-third of subjects were currently using simple analgesics (nonsteroidal anti-inflammatory drugs [NSAIDs], acetaminophen) (29.0%; n = 165) as acute treatment for migraine. A total of 103 subjects (18.1%) were not using any migraine medication (Table 3).

PRO instrument scores of the sample at study entry are presented in Table 4. A majority of subjects (≥75%) had HIT-6[TM] scores (>59) suggesting "severe" headache impact, and mean MSQ scores (<65) that indicated presence of limitations in the areas of daily social and work-related activities. The average PROMIS-PF T-score was 49.2, suggesting overall physical functioning comparable with the U.S. general population norm of 50.[27]

Sample characteristics for the Stage 1 interim sample and the Cohort 2 sample used in Stage 3 to evaluate responsiveness were largely similar to the full study sample. Cohort 2 subjects were starting new preventative medications for migraines or increasing the dose of a current preventative medication upon enrollment into the study. The most common medication in this group was topiramate. Overall, 89% of expected daily diary entries were completed by patients.

Stage 1: Item Analysis, Item Reduction, and Scoring.—The full range of responses (1 to 5) was observed for all items in MFIQ v.1; the median score was 3.0 [moderately difficult/sometimes] for 19 items and 2.0 [slightly/a little difficult/rarely] for the remaining 12 items. Moderate floor effects were observed for items about *difficulty grooming* and *feel like a burden*. No ceiling effects were observed. Average scores on MFIQ items ranged from 2.1 (*item 8: difficulty grooming*) to 3.2 (*item 5: needing to rest or lie down*, and *item 17: presence of loud noises, strong smells, or bright lights*).

EFA was performed on the MFIQ v.1 data at study start (Day 1) to evaluate the domain structure of the items. One-, 2-, 3-, and 4-factor models were evaluated. EFA results suggested good fit for the 1-factor (CFI = 0.953, SRMR = 0.060; RMSEA = 0.123), 2-factor (CFI = 0.969, SRMR = 0.043, RMSEA = 0.104), and 3-factor (CFI = 0.976, SRMR = 0.035, RMSEA = 0.094) models. However, the 4-factor model showed the best fit to the data, as supported by CFI = 0.982 and SRMR = 0.027. In addition, RMSEA (0.086) was only slightly above the conventional cut-off of <0.08 for acceptable fit. Item loadings ranged from 0.391 to 0.892 across the 4 factors. Strong correlations between the factors (0.551-0.560) suggested that these factors also reflected an overarching concept of migraine impact.

IRT analyses supported unidimensionality of the 4 MFIQ domains, and suggested that items generally functioned in an orderly way within the scales and contributed to the measurement of their respective domains. Among physical function items, *ability to move inside home (item 3)* showed evidence of misfit ($P = .003$). When it was omitted, all items had adequate fit ($P \geq .022$). Otherwise, fit statistics did not identify any evidence of substantial item misfit. Usual activities ($P \geq .040$), social function ($P \geq .094$), and emotional function items ($P \geq .037$) all showed adequate fit within their respective domains.

The results of item-level analyses based on a priori criteria were considered along with results from previous qualitative work with patients and input from clinical experts to inform decisions on item selection. The items dropped all had high correlations with other similar items, indicating redundancy, and the most conceptually relevant items were selected in discussion with clinical experts. Five items were dropped, for a total of 26 items in the final MFIQ v.2 (Table 2).

Following item reduction, a separate unidimensional CFA model was estimated for each of the 4 MFIQ v.1 domains. Results showed good fit for physical function (CFI = 0.980; RMSEA = 0.249; WRMR = 1.282; factor loading range: 0.818-0.891), usual activity (CFI = 0.984; RMSEA = 0.134; WRMR = 1.067; factor loading range: 0.797-0.924), social function (CFI = 0.998; RMSEA = 0.119; WRMR = 0.477; factor loading range: 0.851-0.951), and emotional function (CFI = 1.000; RMSEA = 0.000; WRMR = 0.211; factor loading range: 0.839-0.920) domains, and supported the separate 4-factor structure. For 3 of the 4 domains, RMSEA was above the conventional cut-off point of 0.08 for adequate fit; however, RMSEA has been found to be biased to the high side for simple models. The other fit indices, including CFI and WRMR, generally suggested that the separate unidimensional models for each domain provided acceptable fit. These results indicate that all items contributed to the measurement of the underlying concept in their respective domains.

Four domain scores were developed for MFIQ v.2 representing Physical Function (PF; 5 items), Usual Activities (UA; 10 items), Social Function (SF; 5 items), and Emotional Function (EF; 5 items). An additional single item for overall impact on UA was also retained as a global measure. The MFIQ domain summary scores and global score are transformed to a 0-100 scale, with higher scores representing greater impact of migraine.

Stage 2: Reliability and Validity.—Reliability.—All 4 domains showed good to excellent internal consistency at Week 4 (PF: α = 0.91, UA: α = 0.96, SF: α = 0.95, EF: α = 0.94); results were very similar at study start on Day 1 (PF: α = 0.89, UA: α = 0.95, SF: α = 0.94, EF: α = 0.92). Deletion of any single item did not improve reliability in any of the domains.

Good to moderate test–retest reliability was demonstrated among subjects who were stable (based on PGIC "no change in migraine" at Week 4) for MFIQ scores from Day 1 to Week 4 (n = 235). All domain scores and the global UA item score demonstrated good to moderate test–retest reliability (ICCs for PF = 0.57, UA = 0.63, SF = 0.65, EF = 0.71, and global UA = 0.47).

Convergent Validity.—All MFIQ domains had good convergent validity with measures assessing similar constructs (Table 5). Moderate correlations were observed between the number of migraine days and each MFIQ domain score (PF $r = 0.48$; UA $r = 0.48$; SF $r = 0.47$; EF $r = 0.48$) and the global UA item score ($r = 0.48$). Moderate correlations were also observed between the number of headache days and each MFIQ domain score (PF $r = 0.43$; UA $r = 0.41$; SF $r = 0.41$; EF $r = 0.46$) and the global UA item

Table 5.—Construct Validity: Spearman Correlations between MFIQ Domain Scores with Ancillary Measures at Week 4 (N = 569)

Measure	Mean (SD)	n	Spearman correlation (*r*): MFIQ Domains at Week 4				
			Domain 1: Physical Function	Domain 2: Usual Activities	Domain 3: Social Function	Domain 4: Emotional Function	Global Item: Overall Impact on Usual Activities
MFIQ Domain Scores: Mean (SD)		545	39.9 (22.6)	30.8 (22.0)	33.2 (24.8)	43.5 (29.2)	35.8 (24.8)
Headache Frequency (Past Week)							
Number of Migraine Days	2.2 (1.9)	532	0.48*	0.48*	0.47*	0.48*	0.48*
Number of Headache Days	2.8 (2.0)	532	0.43*	0.41*	0.41*	0.46*	0.41*
PRO Score (Past 4 Weeks)							
HIT-6™ Score	61.8 (6.0)	247	0.62*	0.64*	0.62*	0.73*	0.55*
MSQ Role Function-Restrictive Domain Score	54.9 (22.7)	247	−0.73*	−0.77*	−0.78*	−0.74*	−0.67*
MSQ Role Function-Preventive Domain Score	69.3 (24.1)	247	−0.72*	−0. 80*	−0.78*	−0.68*	−0.70*
MSQ Emotional Function Domain Score	60.5 (30.8)	247	−0.60*	−0.64*	−0.67*	−0.81*	−0.60*
PRO Score (Current Status)							
PROMIS-PF Score	48.3 (9.2)	248	−0.54*	−0.55*	−0.51*	−0.46*	−0.45*

*$P < .0001$.
HIT-6 = 6-item Headache Impact Test; MFIQ = Migraine Functional Impact Questionnaire; MSQ = Migraine-Specific Quality-of-Life Questionnaire; PRO = patient-reported outcome; PROMIS-PF = Patient-Reported Outcomes Measurement Information System-Physical Function; SD = standard deviation.

score ($r = 0.41$). Moderate to large correlations were observed between MFIQ v.2 scores and the HIT-6™ score, MSQ RF-R, RF-P, and EF domain scores, and PROMIS-PF score (Table 5).

Known-Groups Validity.—MFIQ v.2 domain scores discriminated between groups of subjects known to be clinically different among groups varying by number of migraine days and headache days ($P < .001$). A pattern was observed of higher MFIQ v.2 domain scores corresponding to groups with greater numbers of migraine days during the 4 weeks, indicating that higher impact was related to having more days with migraine (Fig. 1). MFIQ domain scores also increased with greater impact of headache based on HIT-6™ (Fig. 2). Differences in mean MFIQ UA, PF, SF, and EF scores and HIT-6™ groups were identified for all pairwise comparisons across domains, except for minimal/mild vs moderate levels of impact which had few subjects in these groups. MFIQ domain scores also differentiated by

groups with varying impact on quality of life based on MSQ RF-R, RF-P, and EF domain scores ($P < .0001$ for all MFIQ domains and the global UA item).

All mean MFIQ domain scores increased with higher levels of patient-reported interference of migraine with daily activities. Nearly, all pairwise comparisons of the MFIQ UA, PF, SF, and EF scores and the global UA item with the conceptually similar migraine interference level were nominally significant ($P < .001$).

Stage 3: Responsiveness.—Comparisons of change from Week 4 or 8 to Week 16 in MFIQ domain scores supported the responsiveness of MFIQ (Table 6). Subjects in the study sample with at least 50% reduction in monthly migraine days from Week 4 to Week 16 had significantly more change over time in MFIQ scores than those with less than a 50% reduction, in all domains and the global item ($P < .0001$; mean differences by responder status ranged from −13.00 to −17.18; Fig. 3).
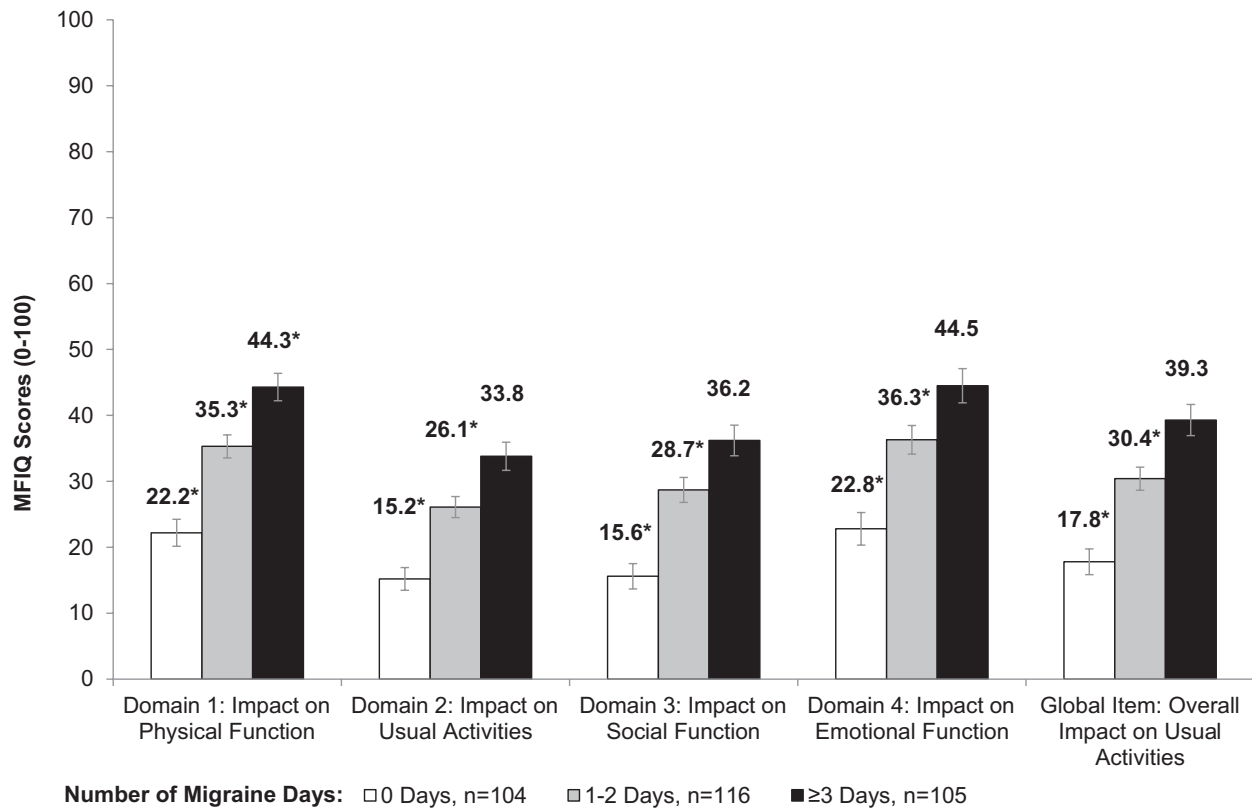
Fig. 1.—Known-Groups Validity: Migraine Functional Impact Questionnaire Domain Scores by Number of Migraine Days at Week 4.

Subjects with a reduction in monthly headache days (≥50%) also showed more improvement in MFIQ scores over time in all domains ($P < .01$) and for the global item. MFIQ change scores were greater for responders (vs nonresponders) based on at least 50% reduction in migraine interference with daily activities (all domains) ($P < .01$).

Large differences in change over time in MFIQ domain scores with change in PGIC and other PROs (HIT-6[TM] and MSQ) also supported the ability of MFIQ to detect changes in migraine impact ($P < .05$). MFIQ domain change comparisons to PGIC responder groups were significant for Improved vs No Change and Improved vs Worsened ($P < .001$), but not for No Change vs Worsened, for all domains. The differences in change over time on the global item score was only significant for the Improved vs Worsened comparison ($P < .01$). Responders based on HIT-6[TM]

(Fig. 4), MSQ domains, and PROMIS-PF showed more improvement on the MFIQ in all domains and for the global UA item, compared with nonresponders ($P < .001$; Table 6).

DISCUSSION

The final 26-item MFIQ v.2 assesses migraine-related functional impacts in 4 domains: physical function, usual activities, emotional function, and social function. The MFIQ was developed with input from patients and feedback from clinical experts and has demonstrated reliability and validity. As expected, the MFIQ was moderately correlated with clinical and PRO measures assessing related constructs. It is responsive to change over time in migraine frequency, interference with daily activities, and PRO assessments. The MFIQ offers comprehensive concept coverage of the impacts of migraine that are most
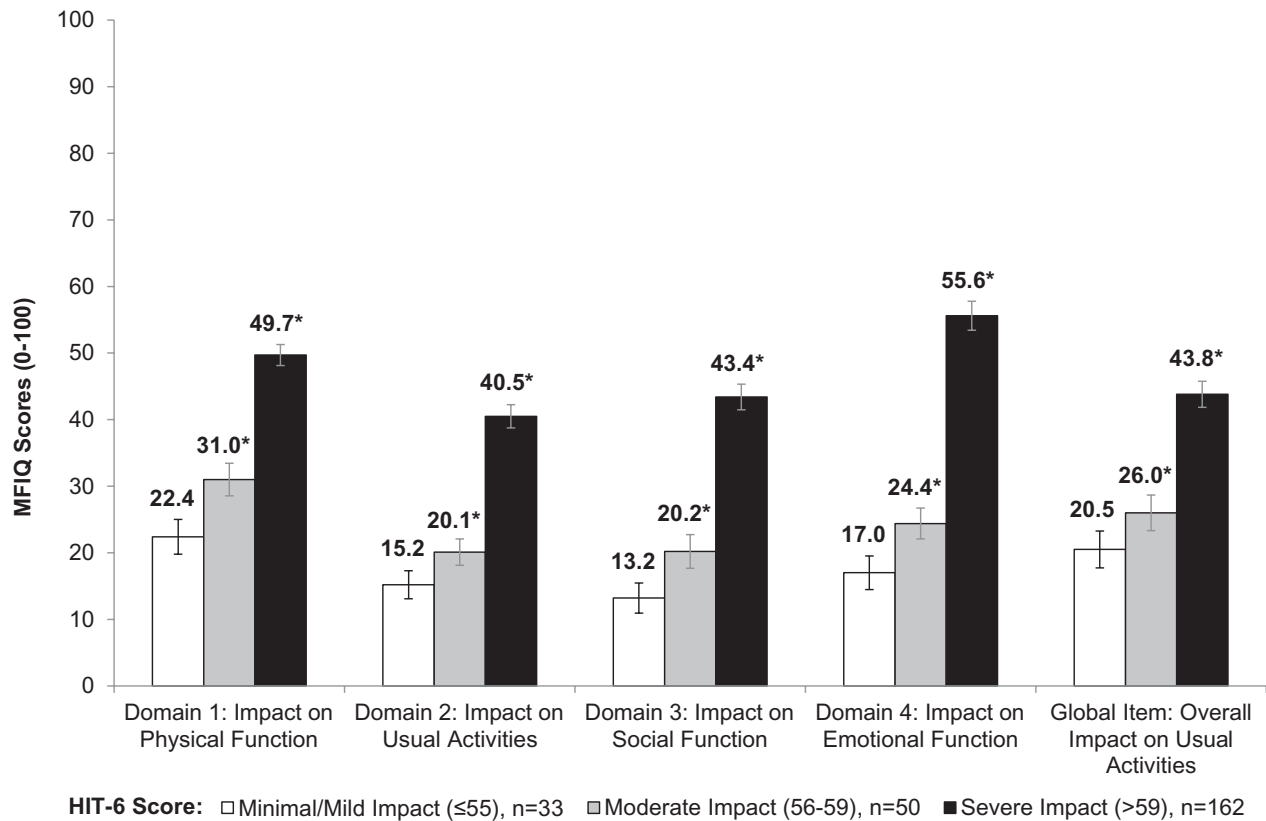
Fig. 2.—Known-Groups Validity: Migraine Functional Impact Questionnaire Domain Scores by HIT-6[TM] Scores at Week 4.

relevant to adults with the condition over a 7-day recall period.

When evaluating interventions intended to treat frequent migraines, a migraine-specific measure such as the MFIQ may be preferred over measures that were developed to assess the impact of headaches in general. Unlike some other PRO measures like the MIDAS and HIT-6[TM], the MFIQ offers separate domain scores enabling evaluating changes in the impact of migraine on both acts and tasks related to functioning and activities. The measure is not limited to solely physical or role impacts, as social and emotional impacts are also captured by the items. However, it was not designed to measure absenteeism from work or other areas of life, so in this regard, it may capture slightly different concepts than measures such as MIDAS or Work Productivity and Activity Impairment Questionnaire (WPAI).[37] Changes in the frequency

and level of difficulty on these psychosocial impacts can be measured by the MFIQ. A recall period of the past 7 days used in the MFIQ enables capture of the experiences of adults experiencing frequent migraines, without the potential risk of recall bias associated with longer recall periods. The instrument may also be sensitive to impact that is not directly related to a migraine day in individuals with less frequent migraines. MFIQ scores suggested that some impairment is present even in subjects with no migraine days during week 4. These impairments in the absence of migraine could potentially be prodrome or postdrome impacts experienced between migraine attacks.

The MFIQ is able to capture the week-to-week variability that is an important aspect of the frequent and episodic nature of migraine[38] and naturally occurring fluctuations in the disease. Although some items in the MFIQ reference activities that may not

Table 6.—Responsiveness: Migraine Functional Impact Questionnaire Mean Change Scores

		MFIQ Domain Score: Mean Change (SD)				
	N	Physical Function	Usual Activities	Social Function	Emotional Function	Global Item: Impact on Usual Activities
Reduction in Monthly Migraine Days†, Mean Difference		−14.51***	−13.00***	−14.37***	−17.18***	−16.54***
Responder	113	−12.8 (2.17)	−9.63 (1.97)	−12.5 (2.22)	−16.4 (2.67)	−13.8 (2.77)
Nonresponder	154	1.72 (1.95)	3.37 (1.77)	1.84 (2.00)	0.73 (2.40)	2.75 (2.51)
Reduction in Monthly Headache Days†, Mean Difference		−13.27***	−11.75***	−12.71***	−16.64***	−11.77**
Responder	106	−13.1 (2.34)	−9.83 (2.13)	−12.6 (2.40)	−17.4 (2.85)	−12.0 (3.04)
Nonresponder	161	0.16 (1.90)	1.92 (1.73)	0.15 (1.95)	−0.80 (2.32)	−0.19 (2.48)
Reduction in Migraine Interference with Daily Activities†, Mean Difference		−13.20***	−11.28***	−12.14***	−13.11***	−9.63*
Responder	111	−12.0 (2.29)	−8.18 (2.10)	−10.9 (2.36)	−13.6 (2.81)	−9.22 (3.00)
Nonresponder	149	1.25 (2.01)	3.11 (1.84)	1.21 (2.08)	−0.48 (2.46)	0.41 (2.65)
HIT-6™ Score‡, Mean Difference		−15.54***	−11.61***	−13.90***	−17.87***	−13.21***
Responder	63	−15.7 (2.47)	−11.4 (2.22)	−14.5 (2.45)	−17.9 (2.97)	−12.9 (2.99)
Nonresponder	207	−0.21 (1.70)	0.17 (1.52)	−0.57 (1.68)	−0.04 (2.04)	0.29 (2.05)
PGIC§, *F*-value, *P*-value for Responder Status¶		26.3	20.6	18.8	25.9	5.85
		1***, 2***, 3 NS	1**, 2***, 3 NS	1**, 2***, 3 NS	1***, 2***, 3 NS	1 NS, 2*, 3 NS
Improved	188	−10.4 (1.83)	−7.13 (1.69)	−9.32 (1.95)	−14.4 (2.22)	−8.13 (2.43)
No change	49	2.75 (2.78)	3.65 (2.58)	3.81 (2.97)	5.92 (3.38)	0.97 (3.70)
Worsened	37	11.93 (3.26)	12.10 (3.00)	7.95 (3.46)	12.45 (3.92)	8.60 (4.51)
MSQ Role Function-Restrictive Domain Score††, Mean Difference		−11.29***	−10.46***	−9.88***	−15.02***	−9.32**
Responder	114	−10.4 (2.03)	−8.51 (1.78)	−9.57 (2.02)	−12.9 (2.40)	−8.23 (2.42)
Nonresponder	155	0.89 (1.88)	1.95 (1.64)	0.30 (1.86)	2.15 (2.22)	1.09 (2.25)
MSQ Role Function-Preventive Domain Score††, Mean Difference		−8.71**	−9.75***	−10.05***	−12.92***	−10.45**
Responder	105	−9.96 (2.29)	−9.26 (1.99)	−10.9 (2.23)	−13.2 (2.69)	−10.1 (2.63)
Nonresponder	164	−1.26 (1.81)	0.49 (1.57)	−0.82 (1.75)	−0.26 (2.14)	0.34 (2.09)
MSQ Emotional Function Domain Score‡‡, Mean Difference		−11.15***	−11.74***	−12.54***	−17.41***	−12.13***
Responder	76	−12.4 (2.46)	−11.5 (2.13)	−13.5 (2.38)	−17.5 (2.86)	−12.1 (2.88)
Nonresponder	193	−1.29 (1.74)	0.27 (1.50)	−0.92 (1.68)	−0.12 (2.02)	0.02 (2.04)
PROMIS-PF Score§§, Mean Difference		−7.84**	−9.30***	−10.54***	−10.80**	−10.91***
Responder	99	−9.46 (2.33)	−9.08 (2.04)	11.2 (2.29)	−11.8 (2.78)	−10.5 (2.71)
Nonresponder	173	−1.62 (1.79)	0.22 (1.56)	−0.66 (1.75)	−1.00 (2.14)	0.43 (2.08)

*$P < .01$.

**$P < .001$.

***$P < .0001$.

†Week 4 to Week 16; Monthly Migraine Day, Monthly Headache Day, or Migraine Interference with Daily Activities Responder ≥50% reduction; Nonresponder <50% reduction.

‡Week 8 to Week 16; HIT-6 Responder ≥5-point reduction; Nonresponder <5-point reduction.

§Week 4 to Week 16; PGIC Improved defined as PGIC of "a little better," "moderately better," or "very much better" at Week 16; no change defined as PGIC of "no change" at Week 16; worsened defined as PGIC of "a little worse," "moderately worse," or "very much worse" at Week 16.

¶1: Improved vs No Change, 2: Improved vs Worsened, 3: No Change vs Worsened.

††Week 8 to Week 16; MSQ RF-R or RF-P Responder ≥5-point increase; Nonresponder <5-point increase.

‡‡Week 8 to Week 16; MSQ EF Responder ≥8-point increase; Nonresponder <8-point increase.

§§Week 8 to Week 16; PROMIS-PF Responder ≥2-point increase; Nonresponder <2-point increase.

EF = Emotional Function; HIT-6 = 6-item Headache Impact Test; MFIQ = Migraine Functional Impact Questionnaire; MSQ = Migraine-Specific Quality-of-Life Questionnaire; NS = not significant; PGIC = Patients' Global Impression of Change; PROMIS-PF = Patient-Reported Outcomes Measurement Information System-Physical Function; RF-P = Role Function-Preventive; RF-R = Role Function-Restrictive; SD = standard deviation.
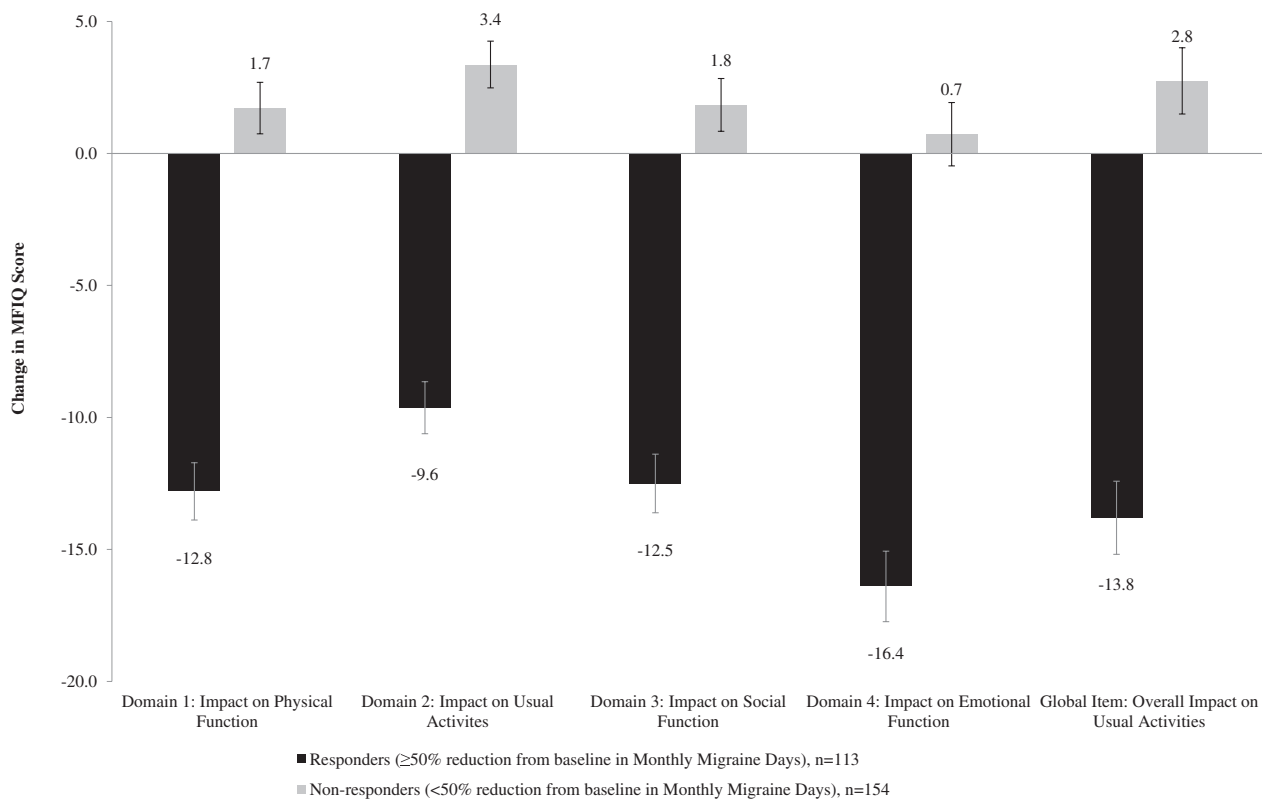
Fig. 3.—Responsiveness: Migraine Functional Impact Questionnaire Domain Change Scores by Responder Status Based on Number of Migraine Days.

occur daily (ie, work or study-related activities), patients have stated that these impacts are relevant and important to their migraine experience. Therefore, for those items that describe concepts that may occur less frequently, the response scale includes a "*does not apply*" option that can be used in cases where the activity was not relevant to a respondent during a particular week. If the MFIQ is administered every week, it may be feasible to average across multiple time points (eg, 4 weeks of assessments) to generate scores that are representative across a longer duration. However, the psychometric properties of multi-week scores have not been evaluated and would require additional analyses to establish their reliability and validity.

The MFIQ is being used as an outcome measure in clinical trials and observational research. It may also be a useful tool for integration into clinical practice. It could be used for initial and ongoing assessments in clinical contexts to monitor outcomes and to enhance communication between patients and healthcare professionals for the management of migraine.

Requests for copies of the instrument may be addressed to the corresponding author.

LIMITATIONS

This study had a few limitations of note. Inclusion of subjects from the Stage 1 sample in the Stage 2 analysis may impact results derived from the subsequent psychometric analyses. However, the total sample included 305 additional subjects beyond those included in the interim sample, so any impacts on the results are considered relatively minimal. The ancillary measures included in the observational study allowed evaluation of convergent validity, but not discriminant validity. Recruitment of the patient sample for the observational study was intentionally limited to match a typical clinical trial sample and was limited to U.S. English. Subjects had to have at least 4 headache days per month to be enrolled in the study, which may also limit generalizability of this sample to the general population of migraine patients.
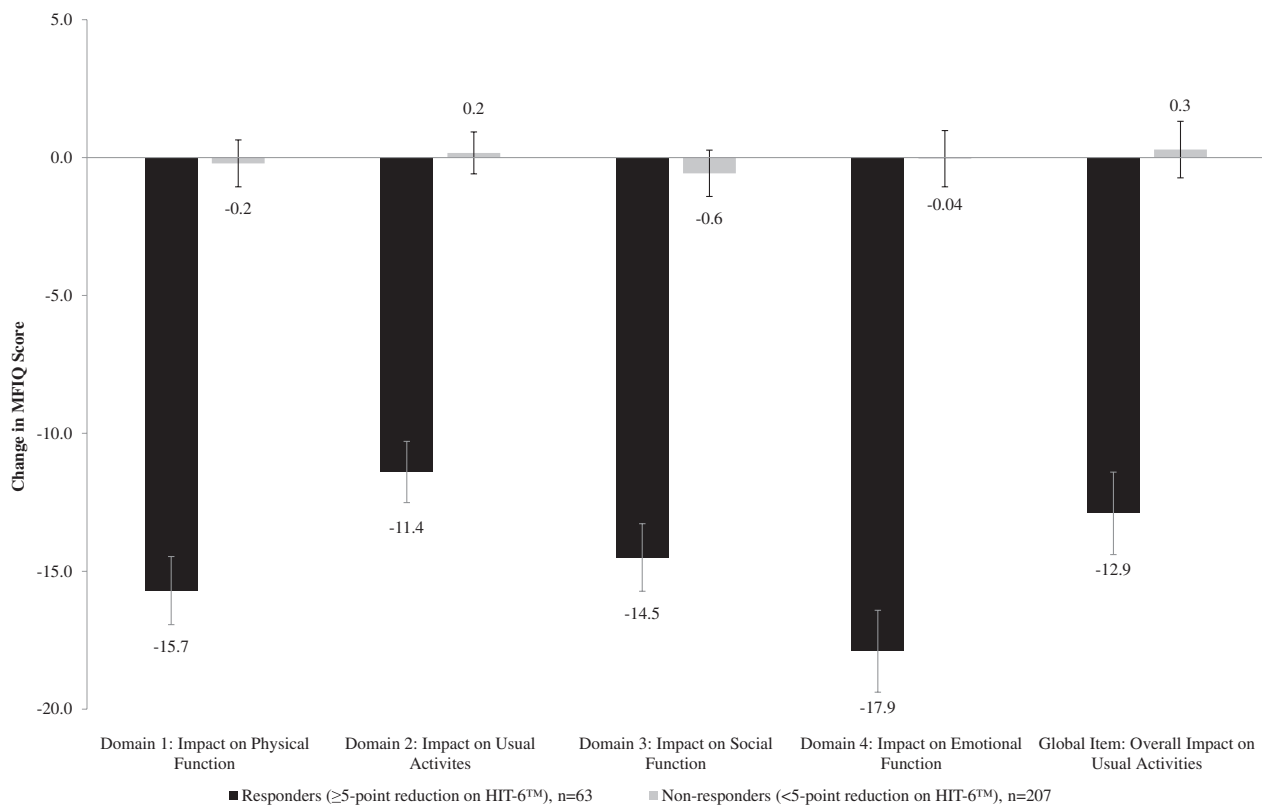
Fig. 4.—Responsiveness: Migraine Functional Impact Questionnaire Domain Change Scores by Responder Status Based on HIT-6™ Score.

CONCLUSIONS

The MFIQ reflects the multifaceted impacts of migraine on functioning that are most relevant to adults with migraine. This study has demonstrated that the MFIQ has robust psychometric properties and is sensitive to change over time. MFIQ scores can be used to assess and track the impact of migraine, evaluate therapeutic targets, identify gaps for intervention, and enhance dialog between patients and healthcare providers about specific impacts of migraine that need to be addressed. It may also be used to evaluate real-world outcomes of interventions for migraine in research and clinical practice settings. Future analyses will be conducted to develop responder definitions for interpreting clinically meaningful within-subject change over time.

STATEMENT OF AUTHORSHIP

Category 1

(a) Conception and Design

Ariane K. Kawata, Asha Hareendran

(b) Acquisition of Data

Shannon Shaffer, Sally Mannix

(c) Analysis and Interpretation of Data

Ariane K. Kawata, Martha Bayliss, Dawn C. Buse, Andrew Thach, Pooja Desai, Daniel D. Mikol, Brian Ortmeier, Asha Hareendran

Category 2

(a) Drafting the Manuscript

Ariane K. Kawata, Shannon Shaffer, Asha Hareendran

(b) Revising It for Intellectual Content

Ariane K. Kawata, Shannon Shaffer, Sally Mannix, Martha Bayliss, Dawn C. Buse, Andrew Thach, Pooja Desai, Daniel D. Mikol, Brian Ortmeier, Asha Hareendran

Category 3

(a) Final Approval of the Completed Manuscript

Ariane K. Kawata, Shannon Shaffer, Sally Mannix, Martha Bayliss, Dawn C. Buse, Andrew Thach, Pooja Desai, Daniel D. Mikol, Brian Ortmeier, Asha Hareendran

REFERENCES

1. Mannix S, Skalicky A, Buse DC, et al. Measuring the impact of migraine for evaluating outcomes of preventive treatments for migraine headaches. *Health Qual Life Outcomes*. 2016;14:143.

2. Buse DC, Scher AI, Dodick DW, et al. Impact of migraine on the family: Perspectives of people with migraine and their spouse/domestic partner in the CaMEO Study. *Mayo Clin Proc*. 2016;91: 596-611.

3. D'Amico D, Grazzi L, Curone M, et al. Difficulties in work activities and the pervasive effect over disability in patients with episodic and chronic migraine. *Neurol Sci*. 2015;36(Suppl. 1):9-11.

4. Raggi A, Giovannetti AM, Quintas R, et al. A systematic review of the psychosocial difficulties relevant to patients with migraine. *J Headache Pain*. 2012; 13:595-606.

5. Blumenfeld AM, Varon SF, Wilcox TK, et al. Disability, HRQoL and resource use among chronic and episodic migraineurs: Results from the International Burden of Migraine Study (IBMS). *Cephalalgia*. 2011;31:301-315.

6. Silberstein SD, Holland S, Freitag F, Dodick DW, Argoff C, Ashman E. Evidence-based guideline update: Pharmacologic treatment for episodic migraine prevention in adults: Report of the Quality Standards Subcommittee of the American Academy of Neurology and the American Headache Society. *Neurology*. 2012;78:1337–1345.

7. Ramadan NM, Silberstein SD, Feitag FG, Gibert TT, Frishberg BM. Evidence-based guidelines for migraine headache in the primary care setting: Pharmacological management for prevention of migraine. *Neurology*. 2012:1-55.

8. Silberstein SD. Practice parameter: Evidence-based guidelines for migraine headache (an evidence-based review): Report of the Quality Standards Subcommittee of the American Academy of Neurology. *Neurology*. 2000;55:754-762.

9. Food Drug Administration (FDA). Guidance for industry patient-reported outcome measures: Use in medical product development to support labeling claims. *Fed Regist*. 2009;74:65132-65133.

10. Kosinski M, Bayliss MS, Bjorner JB, et al. A six-item short-form survey for measuring headache impact: the HIT-6. *Qual Life Res*. 2003;12:963–974.

11. Stewart WF, Lipton RB, Dowson AJ, Sawyer J. Development and testing of the Migraine Disability Assessment (MIDAS) Questionnaire to assess headache-related disability. *Neurology*. 2001;56:S20-S28.

12. Jhingran P, Osterhaus JT, Miller DW, Lee JT, Kirchdoerfer L. Development and validation of the Migraine-Specific Quality of Life Questionnaire. *Headache*. 1998;38:295-302.

13. Badley EM. Enhancing the conceptual clarity of the activity and participation components of the international classification of functioning, disability, and health. *Soc Sci Med*. 2008;66:2335-2345.

14. Hareendran A, Skalicky A, Mannix S, et al. Development of a new tool for evaluating the benefit of preventive treatments for migraine on functional outcomes – The Migraine Functional Impact Questionnaire (MFIQ). *Headache*. 2018; 58: 1612-1628.

15. Norquist JM, Girman C, Fehnel S, DeMuro-Mercon C, Santanello N. Choice of recall period for patient-reported outcome (PRO) measures: criteria for consideration. *Qual Life Res*. 2012;21:1013-1020.

16. Hareendran A, Mannix S, Skalicky A, et al. Development and exploration of the content validity of a patient-reported outcome measure to evaluate the impact of migraine- the migraine physical function impact diary (MPFID). *Health Qual Life Outcomes*. 2017;15:224.

17. Kawata AK, Hsieh R, Bender R, et al. Psychometric evaluation of a novel instrument assessing the impact of migraine on physical functioning: The migraine physical function impact diary. *Headache*. 2017;57:1385-1398.

18. Buse DC, Lipton RB, Mikol DD, et al. Reducing the impact of migraine on functioning: Results from the strive trial: A phase 3, randomized, double-blind study of Erenumab in subjects with episodic migraine. In: *18th Congress of the International Headache Society*. Vancouver, British Columbia; 2017.

19. Kawata AK, Shah N, Poon J-L, et al. Treatment patterns and healthcare resource use in migraine

patients newly initiating a preventive treatment: Interim results from the Assessment of TolerabiliTy and Effectiveness in MigrAINeurs using Preventive Treatment (ATTAIN) Study. In: *Academy of Managed Care Pharmacy (AMCP) Nexus 2017*. Dallas, TX; 2017.

20. Kawata AK, Shah N, Poon J-L, et al. Characteristics of patients newly initiating a preventive treatment for migraine: Baseline data from the Assessment of TolerabiliTy and Effectiveness in MigrAINeurs using Preventive Treatment (ATTAIN) study. In: *18th Congress of the International Headache Society (IHC)*. Vancouver, British Columbia; 2017.

21. Tassorelli C, Diener HC, Dodick DW, et al. Guidelines of the International Headache Society for controlled trials of preventive treatment of chronic migraine in adults. *Cephalalgia*. 2018;38: 815-832.

22. Patrick DL, Burke LB, Gwaltney CJ, et al. Content validity–establishing and reporting the evidence in newly developed patient-reported outcomes (PRO) instruments for medical product evaluation: ISPOR PRO Good Research Practices Task Force report: part 2–assessing respondent understanding. *Value Health*. 2011;14:978-988.

23. Nunnally JC, Bernstein IH. *Psychometric Theory*, 3rd ed. New York: McGraw-Hill; 1994.

24. Gorsuch R. *Factor Analysis*, 2nd ed. Hillsdale, NJ: L. Erlbaum Associates; 1983.

25. FDA Center for Drug Evaluation and Research Department of Health and Human Services. NDA Application No. 20-505/S022; 2004.

26. Dowson AJ. Assessing the impact of migraine. *Curr Med Res Opin*. 2001;17:298-309.

27. Cella D, Riley W, Stone A, et al. The Patient-Reported Outcomes Measurement Information System (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005–2008. *J Clin Epidemiol*. 2010;63:1179-1194.

28. Jhingran P, Davis SM, LaVange LM, Miller DW, Helms RW. MSQ: Migraine-Specific Quality-of-Life Questionnaire. Further investigation of the factor structure. *PharmacoEconomics*. 1998;13:707-717.

29. Muthén LK, Muthén B. *Mplus User's Guide*, 3rd ed. Los Angeles, CA: Muthén & Muthén; 1998-2004.

30. Cai L, Thissen D, du Toit S. *IRTPRO for Windows [Computer Software]*. Lincolnwood, IL: Scientific Software International; 2011.

31. Hu L, Bentler PM. Cutoff criterion for fit indices in covariance structure analysis: Conventional criteria versus new alternatives. *Struct Equ Modeling*. 1999;6:1-55.

32. Browne MW, Cudeck R. Alternative ways of assessing model fit. In: Bollen KA, Long JS, eds. *Testing Structural Equation Models*. Beverly Hills, CA: Sage; 1993:136-162.

33. Yu C. *Evaluating cutoff criteria of model fit indices for latent variable models with binary and continuous outcomes. Doctoral dissertation*. Los Angeles: University of California; 2002.

34. Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika*. 1951;163:297-334.

35. Hinkle D, Wiersma W, Jurs S. *Applied Statistics for the Behavioral Sciences*, 5th ed. Boston, MA: Houghton Mifflin; 2003.

36. Hays RD, Revicki DA. Reliability and validity, including responsiveness. In: Fayers PM, Hays RD, eds. *Assessing Quality of Life in Clinical Trials*. New York: Oxford University Press; 2005:25-39.

37. Reilly MC, Zbrozek AS, Dukes EM. The validity and reproducibility of a work productivity and activity impairment instrument. *PharmacoEconomics*. 1993; 4:353-365.

38. Serrano D, Lipton RB, Scher AI, et al. Fluctuations in episodic and chronic migraine status over the course of 1 year: Implications for diagnosis, treatment and clinical trial design. *J Headache Pain*. 2017; 18:101.